



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/49, C07K 14/16	A2	(11) International Publication Number: WO 00/39303 (43) International Publication Date: 6 July 2000 (06.07.00)
<p>(21) International Application Number: PCT/US99/31272</p> <p>(22) International Filing Date: 30 December 1999 (30.12.99)</p> <p>(30) Priority Data: 60/114,495 31 December 1998 (31.12.98) US 60/156,670 29 September 1999 (29.09.99) US</p> <p>(71) Applicant: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).</p> <p>(72) Inventors: BARNETT, Susan; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US). HARTOG, Karin; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US). MARTIN, Eric; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US).</p> <p>(74) Agents: DOLLARD, Anne, S.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US) et al.</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>Without international search report and to be republished upon receipt of that report.</i></p>
<p>(54) Title: MODIFIED HIV ENV POLYPEPTIDES</p> <p>(57) Abstract</p> <p>Polynucleotide encoding modified HIV Env polypeptides are disclosed. The Env polypeptides are modified so as to expose at least part of the CD4 binding region. Methods of diagnosis, treatment and prevention using the polynucleotides and polypeptides are also provided.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

MODIFIED HIV ENV POLYPEPTIDESTechnical Field

5 The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native β -sheet configurations has been modified. The invention also pertains to methods
10 of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

Background of the Invention

15 The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) *Science* 220:868-871; Gallo et al. (1984) *Science* 224:500-503; Levy et al., (1984) *Science* 225:840-842; Siegal et al., (1981) *N. Engl. J. Med.* 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous
20 system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) *Science* 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) *Nature* 313:277-284; Sanchez-Pescador et al., (1985) *Science* 227:484-492), including three structural genes -- gag, pol and env -- that are common to all
25 retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) *Nature* 326:662-669; Chakrabarti et al., (1987) *Nature*

30 The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded β -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four β -strand structures (β -3, β -2, β -21, β -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the β -sheet conformation of the bridging sheet region is the β -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The β -strand conformation is only stable when incorporated into a β -sheet, where hydrogen bonds with close to optimal geometry are formed between the peptide groups on adjacent β -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of β -sheets, see, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); and A.L. Lehninger, Biochemistry (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (e.g. a chemokine receptor) and subsequent entry of the virus into the cell. (Wyatt, R., et al. (1998) *Nature* 393:705-711; Kwong, P., et al.(1998) *Nature* 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* 86:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* 62:2107-2144; Putney, S., et al. (1986) *Science* 234:1392-1395; Rushe, J. R., et al. (1988) *Proc. Nat. Acad. Sci. USA* 85: 3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA* 83:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* 62:2622-2628; Palker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* 220:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) 316:72-74; Weis, R., et al. (1985) *Nature* (London) 316:69-72; Weis, R., et al. (1986) *Nature* (London) 324:572-575).

Recent work done by Stamatatos et al (1998) *AIDS Res Hum Retroviruses* 14(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1_{SF162} virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-I_{IIIb} also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) *J Virol.* 69:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) *AIDS Res. and Human Retroviruses*. 14:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* 137:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* 67:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have a protective effect.

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans (1999) *AIDS Res. Hum. Ret.* 15(8):689-698; Bolognesi, D., P., et al. (1994) *Ann. Int. Med.* 8:603-611; Haynes, B., F., et al. (1996) *Science* ;271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) *Science* 255:456-459; Burton, D., R. and Montefiori, D. (1997) *AIDS* 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine co-receptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) *Virology*; 204:297-303). These viruses are generally much more resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 co-receptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol.* 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

Summary of the Invention

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the 4- β antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left
5 to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (*e.g.*, gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one modified (*e.g.*, deleted or replaced)
10 amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (*e.g.*, V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to
15 HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (*e.g.*, amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

20 In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (*e.g.*, deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (*e.g.*, amino acid residues 427 to 429 relative to HXB-2). These polypeptides may have modifications (*e.g.*, a deletion
25 or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes a vaccine composition comprising any of the
30 polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising

(a) administering a first composition comprising any of the polynucleotides described above in a priming step and

(b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV-1_{HXB-2} Env gp120 polypeptide, as determined by crystallography studies.

Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1_{HXB-2} Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5 Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

10 Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204 (SEQ ID NO:3).

Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201 (SEQ ID NO:4).

Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B (SEQ ID NO:5).

15 Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200 (SEQ ID NO:6).

Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199 (SEQ ID NO:7).

20 Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202 (SEQ ID NO:8).

Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431 (SEQ ID NO:9).

Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431 (SEQ ID NO:10).

25 Figure 14 depicts the nucleotide sequence of the construct designated Arg426-Gly431B (SEQ ID NO:11).

Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432 (SEQ ID NO:12).

30 Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432 (SEQ ID NO:13).

Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433 (SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

5 Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

10 Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199; Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200; Asn425-Lys432 (SEQ ID NO:21).

15 Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201; Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

20 Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202; Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEQ ID NO:25).

25 Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

Detailed Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully
30 in the literature. See, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. HIV Protocols in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., Molecular Cloning: A

Laboratory Manual (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. Current Protocols in Molecular Biology, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, Reviews in Computational Chemistry, volumes 1-present (Wiley-VCH, New York, New York, 1999).

- 5 It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

10 **Definitions**

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

- The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.
- 15

- A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2-(methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-1-enyl)cysteine, homoserine, ornithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, cyclohexylalanine, mimosine, pipercolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.
- 20
- 25

- By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.
- 30

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

By "Env polypeptide" is meant a molecule derived from an envelope protein, preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1_{HXB-2} (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel β -strands (β -2, β -3, β -20 and β -21) that form a β -sheet. Extruding from one pair of the β -strands (β -2 and β -3) are two loops, V1 and V2. The β -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while β -3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, e.g., Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatatos et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of β -strands (β -20 and β -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2, β -20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while β -21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue. The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV_{IIIb}, HIV_{SF2}, HIV-1_{SF162}, HIV-1_{SF170}, HIV_{LAV}, HIV_{LAI}, HIV_{MN}, HIV-1_{CM235}, HIV-1_{US4}, other HIV-1 strains from diverse subtypes (e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2_{UC1} and HIV-2_{UC2}), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify β -sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (*e.g.*, "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (*i.e.*, the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

Thus, a "modified Env polypeptide" is an Env polypeptide (*e.g.*, gp120 as defined above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (*e.g.*, gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (*e.g.*, correct geometry). Thus, modifications to the β -20 and β -21 regions (between about amino acid residues 420 and 435 relative to HXB-2) are preferred. Additionally, modifications to the β -2 and β -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (*e.g.*, truncations) to the V1 and V2 loop regions may also be made. Although not all possible β -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, *e.g.*, polyacrylamide gel electrophoresis and the like, and immunological techniques such as Western blotting and immunoprecipitation assays as described in, *e.g.*, International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., *Proc. Natl. Acad. Sci. USA* (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., *Molecular Immunology* (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the *Wisconsin Sequence Analysis Package Program Manual, Version 8* (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH

package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: <http://www.ncbi.nlm.gov/cgi-bin/BLAST>.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as β -sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (e.g., using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) *Methods Enzymol.* 266:540-553; Geourjon et al. (1995) *Comput. Applic. Biosci.* 11:681-684; Levin (1997) *Protein Eng.* 10:771-776; and Rost et al. (1993) *J. Molec. Biol.* 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., *supra*; *DNA Cloning, supra*; *Nucleic Acid Hybridization, supra*.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, e.g., a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemilumescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acradimum esters, NADPH, α - β -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

Overview

The present invention concerns modified Env polypeptide molecules (e.g., glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (*e.g.*, an antibody response) to epitopes in or near the CD4 binding site.

Various forms of the different embodiments of the invention, described herein, may be combined.

β -Sheet Conformations

In the present invention, location of the β -sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between β -strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

Polypeptide Production

The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (*e.g.*, gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard techniques and, *e.g.*, restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. *See, e.g.*, Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (*e.g.*, truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. *See, e.g.*, Edge (1981) *Nature* 292:756; Nambair et al. (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311; Stemmer et al. (1995) *Gene* 164:49-53.

Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. *See, e.g.*, Innis et al, (1990) *PCR Applications: Protocols for Functional Genomics*; Zoller and Smith, *Methods Enzymol.* (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. *See, e.g.*, Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage λ (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), YIp5 (*Saccharomyces*), YCp19 (*Saccharomyces*) and bovine papilloma virus (mammalian cells). See, generally, *DNA Cloning: Vols. I & II, supra*; Sambrook *et al., supra*; B. Perbal, *supra*.

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, e.g., Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987). Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta *et al.*, *Mol. Biotech.* (1996) 5:209-221; and Hackland *et al.*, *Arch. Virol.* (1994) 139:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei *et al.*, *J. Virol.* (1993) 67:4017-4026 and Selby *et al.*, *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. *See, e.g.,* U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; *i.e.*, to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. *See, e.g.,* Sambrook *et al., supra; DNA Cloning*, Vols. I and II, *supra; Nucleic Acid Hybridization, supra.*

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (*e.g.*, Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find

use with the present expression constructs. Yeast hosts useful in the present invention include *inter alia*, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guilliermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Insect cells for use
5 with baculovirus expression vectors include, *inter alia*, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described
10 above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA)
15 leader sequence, a γ -interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoabsorbent
20 techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env
25 polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat
30 treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

Following disruption of the cells, cellular debris is removed, generally by
5 centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

For example, one method for obtaining the intracellular Env polypeptides of the
10 present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from *Galanthus nivalis* agglutinin (GNA), *Lens culinaris* agglutinin (LCA or lentil lectin), *Pisum sativum* agglutinin (PSA or pea lectin), *Narcissus pseudonarcissus*
15 agglutinin (NPA) and *Allium ursinum* agglutinin (AUA). The choice of a suitable affinity resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (e.g., gp120) complexes, either with itself or other
20 proteins. Such complexes are readily produced by e.g., co-transfecting host cells with constructs encoding for the Env (e.g., gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in *trans* or *cis*, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively,
25 the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate. Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See,
30 International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-fluorenylmethoxycarbonyl (Fmoc) benzyloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropylloxycarboxy-carbonyl, t-amylloxycarbonyl, isobornyloxycarbonyl, o-bromobenzyloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

5 **Diagnostic and Vaccine Applications**

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive
10 antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as
15 competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other
20 methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidene fluoride; diazotized paper; nylon membranes;
25 activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating
30 selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase, 5 Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of 10 the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies 15 thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the 20 proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

25 The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral 30 isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125--ser125), GM-CSF, IL-

12, γ -interferon, IP-10, MIP1 β and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (*e.g.*, DNA), using viral vectors (*e.g.*, retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (*e.g.*, liposomes, particles coated with nucleic acid or protein). The vaccines may also
5 comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (*e.g.*, a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

10 The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

A carrier is optionally present which is a molecule that does not itself induce the
15 production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env
20 polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion
25 formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y
30 microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

emulsion, and (c) Ribi™ adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); (3) saponin adjuvants, such as Stimulon™ (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

Experimental

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

- 5 Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

EXAMPLE 1**10 A.1. Best-Fit and Homology Searches**

- The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GCI TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands β -3 and β -2 are connected by the V1/V2 loop. Strands β -21 and β -20 are connected by another small loop. The H-bonds at the interface between strands β -2 and β -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the CD4 binding.

- 20 Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

- 30 Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands (β -20 and β -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the β -2 strand (residues 119-123 of HXB-2), the β -3 strand (residues 199-201 of HXB-2) or both β -2 and β -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

As V1/V2 and one or more residues of β -2 and β -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

Rank	Sequence	RMSD	% Homology	Seq Id. No.
Best fit	GLN-VAL-HIS-ASP-GLU	0.154764	68.75	32
2	LYS-GLU-GLY-ASP-LYS	0.15718	81.25	33
9	ARG-SER-GLY-ARG-SER	0.173731	68.75	34
11	THR-LEU-GLY-ASN-SER	0.175554	81.25	35
16	HIS-PHE-GLY-ALA-GLY	0.178772	93.75	36

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

A.3. One Additional Residue Excisions

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33	37
Best hom:	ASP-SER-ASP-GLY-PRO	0.440896	83.33	38

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional residue from β -2 or β -3 excised were also searched. The V1/V2 loop the target to match was: VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

The construct encoding Ala-Pro (e.g., #7) was recommended.

A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the β -20 and β -21 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from β -2, β -3 or at least one residue excised from β -2 and β -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

5

Table 6: Search of V1/V2 loop (Cys119 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

10

It was determined that both constructs would be used.

B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- β antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire β -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (e.g., tertiary structure of β -20 and β -21), as determined by searching the Brookhaven Protein Data Bank.

20

Table 7A is a summary of the truncations of the variable regions 1 and 2 recommended for this study, as determined in Example 1.A. above.

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120-PRO-GLY-ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having then wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number
128	Serine (S)	Thr (T)/114
195	Serine (S)	Asn (N)/188
426	Met (M)	Arg (R)/411

Constructs containing deletions in the β -20 strand, β -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B). Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

Small Loop/ β -20 and β -21 (Modified)	SEQ ID NO	Figure
-TRP427-GLY-GLY431-	9	12
-ARG426-GLY-GLY-GLY431-	10	13
-ARG426-GLY-SER-GLY431B-	11	14
-ARG426-GLY-GLY-ASN-LYS432-	12	15
-ASN425-ALA-PRO-LYS432-	13	16
-ILE424-GLY-GLY-ALA433-	14	17
-ILE423-GLY-GLY-MET434-	15	18
GLN422-GLY-GLY-TYR435-	16	19
-GLN422-ALA-PRO-TYR435B-	17	20

The deletion constructs shown in Tables 7 and 8 for each one of the β -strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

Exemplary constructs for SF162 are shown in the

Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Val120-Ala204	3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201
Val120-Ile201B	5	8	V1/V2: Val120-Pro-Gly-Ile201
Lys121-Val200	6	9	V1/V2: Lys121-Ala-Pro-Val200

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Val120-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Tyr435
Val127-Asn195	25	28	bsm: Val127-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422-Ala-Pro-Tyr435
Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427- Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Trp427-Gly-Gly431
Lys121-Val200- Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 --- Asn425-Ala-Pro-Lys432
Val120-Ile201-Ile424- Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 --- Ile424-Gly-Gly-Ala433
Val120-Ile201B-Ile424- Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 --- Ile424-Gly-Gly-Ala43
Val120-Thr202; Ile424- Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 --- Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 --- Arg426-Gly-Gly-Gly431

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.

The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in
5 addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein
10 vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

Claims:

What is claimed is:

- 5 1. A polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one amino acid deleted or replaced in the region corresponding to residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).
- 10 2. The polynucleotide of claim 1, wherein the region corresponding to residues 124-198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
- 15 3. The polynucleotide of claim 1, wherein at least one amino acid in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
- 20 4. The polynucleotide of claim 2, wherein at least one amino acid of the in the region corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
- 25 5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
- 30 7. The polypeptide of claim 6, wherein one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

5 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

10 10. The polypeptide of claim 6, wherein at least one amino acid residue between about amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.

15 12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptide are truncated.

13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.

20 14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).

25 15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).

16. A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).

30 17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ ID NO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).
19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ ID NO:8).
20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).
21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).
22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).
23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).
24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ ID NO:13).
25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).
26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).
27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).
28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

5 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ ID NO:19).

31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).

10 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).

33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).

15

34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).

35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

20

36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).

25 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).

38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.

30

39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.

5

42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.

10

43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.

15

44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject

20

45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.

25

46. A method of inducing an immune response in a subject comprising
(a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and
(b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.

30

47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

gp120 core structure

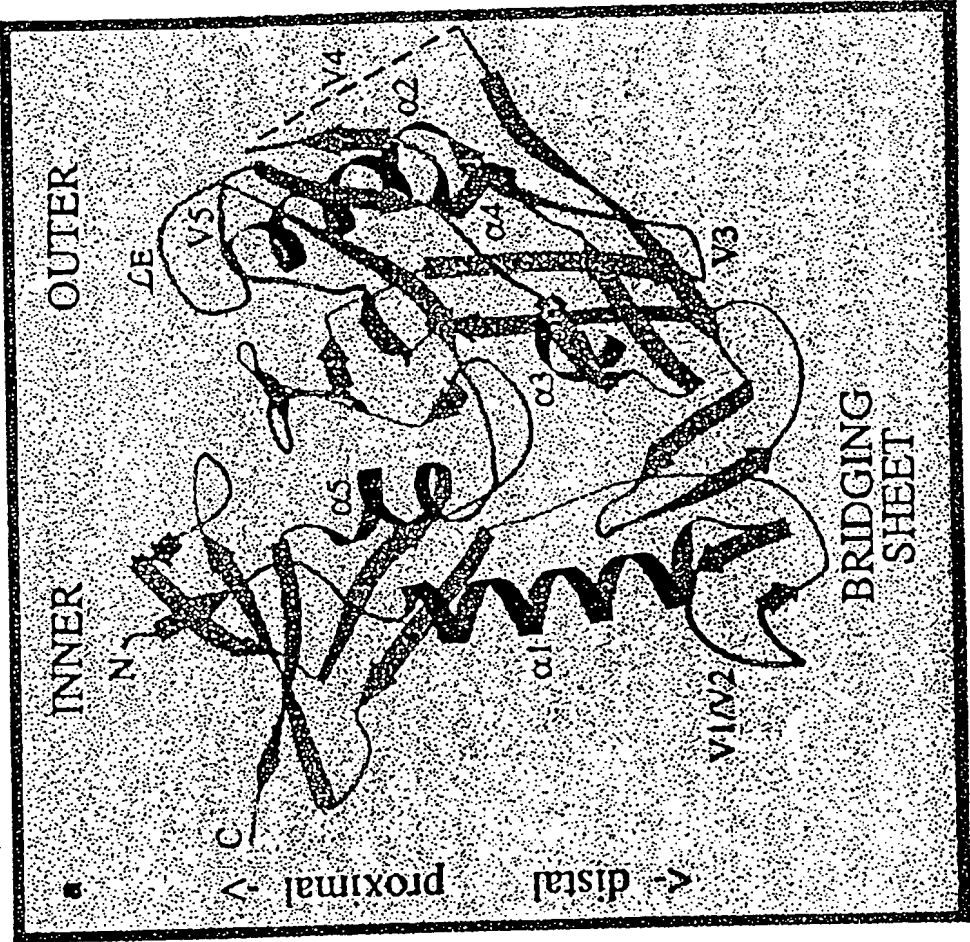


FIG. 1

		1		50
HXB2	(1)	MRVK---EKQHLWRWG	SWRGTLLGLMIC-SATEK	GVVAVK
162	(1)	-----MDAMKRE	LCCVLLCGAFSPS	VEKGVVAVK
SF2	(1)	MRVKGTRRNQHLWRWG	-----TLLGLMIC-SATEK	GVVAVK
CM236	(1)	MRVKETQMNPNLWKWG	-----TLLGLMIC-SASNN	GVVAVK
US4	(1)	--MR---KHCQHLWRG	-----ILLGLMIC-RATTV	GVVAVK
Consensus	(1)	MRVK	YQHLWRWG	TLLGLMIC SATEKLVVTVYYGVVAVK
		51		100
HXB2	(47)	EATLHFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
162	(41)	EATLHFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
SF2	(46)	EATLHFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
CM236	(46)	DADLHFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
US4	(41)	EATLHFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
Consensus	(51)	EATTTLFGASDAKAYDTEV	HNVWATHACVPTDPNPQEVVL	NVTENFNMW
		101		150
HXB2	(97)	KNDVAFQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
162	(91)	KNNVAFQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
SF2	(96)	KNNVAFQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
CM236	(96)	KNNVAFQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
US4	(91)	KNNVAFQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
Consensus	(101)	KNNMVEQMHEDIISLWDQSLKPCVKLT	PLCVTLNCTDL	-----
		151		200
HXB2	(135)	-----KNDTNTN	SSSGRMIEGGEIKNCSFNITTS	IRDKVQKEYALFY
162	(129)	-----KNTNTK	SSNWEMD-RGEIKNCSFNITTS	IRDKVQKEYALFY
SF2	(134)	-----GKTNTN	SSNWKEE-RGEIKNCSFNITTS	IRDKVQKEYALFY
CM236	(135)	-----LTNVNNT	SVSNTIGNITD-RGEIKNCSFNITTS	IRDKVQKEYALFY
US4	(141)	GTNSTSGTNTSTNTS	DSWEKMPGEIKNCSFNITTS	IRDKVQKEYALFY
Consensus	(151)	NATNTNSS	KE M KGEIKNCSFNITTS	IRDKVQKEYALFY
		201		250
HXB2	(178)	KLDVVPIDNDTS	-----YRLINCNTSVITQACPKVSFEPIPIHYCAPAG	
162	(171)	KLDVVPIDNDNTS	-----YRLINCNTSVITQACPKVSFEPIPIHYCAPAG	
SF2	(176)	NLDVVPIDNDAST	ITNYTNRIIHNRATITQACPKVSFEPIPIHYCAPAG	
CM236	(179)	KLDVVPIDNDKTS	-----SEARINNTSVITQACPKVSFEPIPIHYCAPAG	
US4	(191)	KLDVVPIDNDNAS	-----YRLINCNTSVITQACPKVSFEPIPIHYCAPAG	
Consensus	(201)	KLDVVPIDND TS	YRLINCNTSVITQACPKVSFEPIPIHYCAPAG	
		251		300
HXB2	(223)	FADLKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
162	(216)	FADLKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
SF2	(226)	FADLKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
CM236	(226)	FADLKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
US4	(236)	FADLKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
Consensus	(251)	FAILKNDK	FNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAEEVVI	
		301		350
HXB2	(273)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	
162	(266)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	
SF2	(276)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	
CM236	(276)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	
US4	(286)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	
Consensus	(301)	RSENFIDNAKTIIVQLNESVEINCTRPNNNTRKSI	I GPGRAFY TGD	

FIG. 2A

		351		400
HXB2	(323)	M-GNMQAHENISAKNNNTLKKIASKIREQEGNKKUMIEKQSSGGPPEI		
162	(314)	TIIDIRQAHCNISGEKNNNTLKKIVTKQAQEG-NKQFAEKQSSGGPPEI		
SF2	(324)	TIIDIRKAHCNISLAQNNNTLEIVKQIREQEGNKKUMIENQSSGGPPEI		
CM236	(324)	TIIDIRKAYGEENGTKNEVTQVTEKKEHEN-NKUMIEOPPSSGGLLEI		
US4	(334)	TIIDIRQAHCNISKANNTNTLEIIVEKIREQEGNKKUMIENSSGGPPEI		
Consensus	(351)	IIGDIRQAHCNISRAKWNNTL QIV KLREQFGNNKTIIFNQSSGGDPEI		
		401		450
HXB2	(372)	VTISNCGGSEKSTQDENSWFNSTWSIEGSNNTEGSDITTPCRK		
162	(363)	VMHSNCGGSEKSTQDENSW-NN---TIGPNNTNG---TTPCRK		
SF2	(374)	VMHSNCGRSTKSTQDENNWRLN-HLEG---TKGNDITTPCRK		
CM236	(373)	TMHHSNCRSEKSTQDENNCIEN--GIMG--GCNG---TTPCRK		
US4	(384)	VFHSNCGGSEKSTQDENSW--N--IEEVNKTENDITTPCRK		
Consensus	(401)	VMHSFNCGGSEFFYCNTTQLFNSTW N TEG N T G DTIILPCRK		
		↓		
		451		500
HXB2	(422)	QIINMWQKVKAMYPAPISQIRCSNIRGLITRDGG---NSNNITDTEIF		
162	(407)	QIINRWQEVKAMYPAPISQIRCSNIRGLITRDGGK---EISNTT		
SF2	(419)	QIINMWQEVKAMYPAPISQIRCSNIRGLITRDGGT---NNTNT		
CM236	(417)	QIINMWQAGQAMYPAPISGRINVENINGNITRDGG---AINTNT		
US4	(430)	QIINMWQEVKAMYPAPISQIRCSNIRGLITRDGGTNNRNITDTEIF		
Consensus	(451)	QIINMWQEVGKAMYAPPI QIRCSNITGLLLTRDGG NITNDTEIF		
		501		550
HXB2	(469)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI- GA		
162	(455)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI- TE		
SF2	(467)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI- GIV		
CM236	(464)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI- GI		
US4	(480)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI- GI		
Consensus	(501)	RPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQREKRAVGI GA		
		551		600
HXB2	(518)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
162	(504)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
SF2	(517)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
CM236	(513)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
US4	(529)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
Consensus	(551)	MFLGLGAAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQ		
		601		650
HXB2	(568)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		
162	(554)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		
SF2	(567)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		
CM236	(563)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		
US4	(579)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		
Consensus	(601)	LTVWGIKQLQARVLAVERYLKDQQLLGWCSGKLICTTAVPWNASWSNK		

FIG. 2B

		651		700
HXB2	(618)	SLEQANNHTWMEWDRENNYTSLSLHSLIEESQNOQEKNEQEFHSBKKA		
162	(604)	SLEQANNMTWMEWEREDNYTNLIYTLIEESQNOQEKNEQEFHSBKKA		
SF2	(617)	SLEQANNMTWMEWEREDNYTNLIYTLIEESQNOQEKNEQEFHSBKKA		
CM236	(613)	SLEQANNMTWMEWEREDNYTNLIYTLIEESQNOQEKNEQEFHSBKKA		
US4	(629)	SLEQANNMTWMEWEREDNYTNLIYTLIEESQNOQEKNEQEFHSBKKA		
Consensus	(651)	SLEEIWNMTWMEWEREI NYTNLIYTLIEESQNOQEKNEQELLELDKWA		
		701		750
HXB2	(668)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
162	(654)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
SF2	(667)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
CM236	(663)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
US4	(679)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
Consensus	(701)	SLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSF		
		751		800
HXB2	(718)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
162	(704)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
SF2	(717)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
CM236	(713)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
US4	(729)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
Consensus	(751)	QTRLPPRGPDREPEGIEEGGERDRDRSVRLV G LALIWDLLRSLCLFS		
		801		850
HXB2	(768)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
162	(754)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
SF2	(767)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
CM236	(763)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
US4	(779)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
Consensus	(801)	YHRLRDLLIAARIVELLGR-----RGWEALKYWWNLLQYW QELKNS		
		851		900
HXB2	(811)	AVSLLNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		
162	(797)	AVSLFNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		
SF2	(810)	AVSLFNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		
CM236	(813)	AVSLFNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		
US4	(822)	AVSLFNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		
Consensus	(851)	AVSLLNATAIAAEGTDRVIEVAQRAFRAILHIPRRIRQGLER LL		

FIG. 2C

	1	40
Leu122-Ser199	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val127-Asn195	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Ile201B	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Ala204	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Ile201	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Val120-Thr202	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Lys121-Val200	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Consensus	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
	41	80
Leu122-Ser199	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Val127-Asn195	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Val120-Ile201B	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Val120-Ala204	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Val120-Ile201	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Val120-Thr202	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Lys121-Val200	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
Consensus	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAG
	81	120
Leu122-Ser199	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val127-Asn195	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201B	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ala204	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Ile201	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Val120-Thr202	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Lys121-Val200	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTG
	121	160
Leu122-Ser199	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Val127-Asn195	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Val120-Ile201B	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Val120-Ala204	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Val120-Ile201	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Val120-Thr202	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Lys121-Val200	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
Consensus	(121)	CCCGTGTGGAAGGAGGCCACCACCACCTGTTCTGCGCCA
	161	200
Leu122-Ser199	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Val127-Asn195	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Val120-Ile201B	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Val120-Ala204	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Val120-Ile201	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Val120-Thr202	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Lys121-Val200	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
Consensus	(161)	GCGACGCCAAGGCGTACGACACCGAGGTGCACAACGTGTG
	201	240
Leu122-Ser199	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Val127-Asn195	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Val120-Ile201B	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Val120-Ala204	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Val120-Ile201	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Val120-Thr202	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Lys121-Val200	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
Consensus	(201)	GGCCACCCACGCGTGGGTGCCACCGACCCCAACCCCCAG
	241	280
Leu122-Ser199	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAATTCAACATGT
Val127-Asn195	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAATTCAACATGT

FIG. 3A

Val120-Ile201B	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	
Val120-Ala204	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	
Val120-Ile201	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	
Val120-Thr202	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	
Lys121-Val200	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	
Consensus	(241)	GAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGT	281 320
Leu122-Ser199	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Val127-Asn195	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Val120-Ile201B	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Val120-Ala204	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Val120-Ile201	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Val120-Thr202	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Lys121-Val200	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Consensus	(281)	GGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	321 360
Leu122-Ser199	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTG	
Val127-Asn195	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTG	
Val120-Ile201B	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCC----	
Val120-Ala204	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG----	
Val120-Ile201	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG----	
Val120-Thr202	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGG----	
Lys121-Val200	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAG--	
Consensus	(321)	CAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTG	361 400
Leu122-Ser199	(361)	-----GGCAA-----CAGCG	
Val127-Asn195	(361)	ACCCCCCTGTGCGTGGGGGAGGGAAGTGAACACCAGCG	
Val120-Ile201B	(357)	-----CG	
Val120-Ala204	(357)	-----CG	
Val120-Ile201	(357)	-----CG	
Val120-Thr202	(357)	-----CG	
Lys121-Val200	(359)	-----C-----CCCCG	
Consensus	(361)	CG	401 440
Leu122-Ser199	(371)	TGATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Val127-Asn195	(401)	TGATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Val120-Ile201B	(359)	GCATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Val120-Ala204	(357)	----CGCGGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Val120-Ile201	(359)	GCATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Val120-Thr202	(359)	GCGCCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Lys121-Val200	(365)	TGATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Consensus	(401)	ATCAGCCAGGCGCTGCCCAAGGTGAGCTTCGAGCCCAT	441 480
Leu122-Ser199	(411)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Val127-Asn195	(441)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Val120-Ile201B	(399)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Val120-Ala204	(393)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Val120-Ile201	(399)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Val120-Thr202	(399)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Lys121-Val200	(405)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	
Consensus	(441)	CCCCATCCACTACTGCGCGCGCGCGGCTTCGCCATCCTG	481 520
Leu122-Ser199	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Val127-Asn195	(481)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Val120-Ile201B	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Val120-Ala204	(433)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Val120-Ile201	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	

FIG. 3B

Vall120-Thr202	(439)	AAGTGAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA
Lys121-Val200	(445)	AAGTGAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA
Consensus	(481)	AAGTGAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA 521 560
Leu122-Ser199	(491)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Vall127-Asn195	(521)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Vall120-Ile201B	(479)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Vall120-Ala204	(473)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Vall120-Ile201	(479)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Vall120-Thr202	(479)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Lys121-Val200	(485)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC
Consensus	(521)	CCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCGCCC 561 600
Leu122-Ser199	(531)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Vall127-Asn195	(561)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Vall120-Ile201B	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Vall120-Ala204	(513)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Vall120-Ile201	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Vall120-Thr202	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200	(525)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Consensus	(561)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC 601 640
Leu122-Ser199	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Vall127-Asn195	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Vall120-Ile201B	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Vall120-Ala204	(553)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Vall120-Ile201	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Vall120-Thr202	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Lys121-Val200	(565)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA
Consensus	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACA 641 680
Leu122-Ser199	(611)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Vall127-Asn195	(641)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Vall120-Ile201B	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Vall120-Ala204	(593)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Vall120-Ile201	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Vall120-Thr202	(599)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Lys121-Val200	(605)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA
Consensus	(641)	ACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGA 681 720
Leu122-Ser199	(651)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Vall127-Asn195	(681)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Vall120-Ile201B	(639)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Vall120-Ala204	(633)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Vall120-Ile201	(639)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Vall120-Thr202	(639)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Lys121-Val200	(645)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC
Consensus	(681)	GATCAACTGCACCCGCGCCCAACAACAACACCCGCAAGAGC 721 760
Leu122-Ser199	(691)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Vall127-Asn195	(721)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Vall120-Ile201B	(679)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Vall120-Ala204	(673)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Vall120-Ile201	(679)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Vall120-Thr202	(679)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Lys121-Val200	(685)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG
Consensus	(721)	ATCACCATCGGCCCCGGCGCGCCTTCTACGCCACCGGCG

FIG. 3C

	761	800
Leu122-Ser199	(731) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Val127-Asn195	(761) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Val120-Ile201B	(719) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Val120-Ala204	(713) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Val120-Ile201	(719) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Val120-Thr202	(719) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Lys121-Val200	(725) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
Consensus	(761) ACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAG	
	801	840
Leu122-Ser199	(771) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Val127-Asn195	(801) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Val120-Ile201B	(759) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Val120-Ala204	(753) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Val120-Ile201	(759) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Val120-Thr202	(759) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Lys121-Val200	(765) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
Consensus	(801) CGGCGAGAAGTGGAAACAACACCCTGAAGCAGATCGTGACC	
	841	880
Leu122-Ser199	(811) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Val127-Asn195	(841) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Val120-Ile201B	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Val120-Ala204	(793) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Val120-Ile201	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Val120-Thr202	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Lys121-Val200	(805) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
Consensus	(841) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCA	
	881	920
Leu122-Ser199	(851) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Val127-Asn195	(881) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Val120-Ile201B	(839) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Val120-Ala204	(833) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Val120-Ile201	(839) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Val120-Thr202	(839) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Lys121-Val200	(845) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
Consensus	(881) AGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAG	
	921	960
Leu122-Ser199	(891) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val127-Asn195	(921) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val120-Ile201B	(879) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val120-Ala204	(873) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val120-Ile201	(879) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val120-Thr202	(879) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Lys121-Val200	(885) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Consensus	(921) CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACC	
	961	1000
Leu122-Ser199	(931) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val127-Asn195	(961) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val120-Ile201B	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val120-Ala204	(913) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val120-Ile201	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val120-Thr202	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Lys121-Val200	(925) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Consensus	(961) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
	1001	1040
Leu122-Ser199	(971) ACAACACCAACGGCACCATCACCCCTGCCCTGCCGCATCAA	
Val127-Asn195	(1001) ACAACACCAACGGCACCATCACCCCTGCCCTGCCGCATCAA	

FIG. 3D

Val120-Ile201E	(959)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA
Val120-Ala204	(953)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA
Val120-Ile201	(959)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA
Val120-Thr202	(959)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA
Lys121-Val200	(965)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA
Consensus	(1001)	ACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAA 1041 1080
Leu122-Ser199	(1011)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val127-Asn195	(1041)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201B	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ala204	(993)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Ile201	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Val120-Thr202	(999)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Lys121-Val200	(1005)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
Consensus	(1041)	GCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG 1081 1120
Leu122-Ser199	(1051)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val127-Asn195	(1081)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201B	(1039)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ala204	(1033)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ile201	(1039)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Thr202	(1039)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Lys121-Val200	(1045)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Consensus	(1081)	TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA 1121 1160
Leu122-Ser199	(1091)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val127-Asn195	(1121)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ile201B	(1079)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ala204	(1073)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ile201	(1079)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Thr202	(1079)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Lys121-Val200	(1085)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Consensus	(1121)	ACATCACC GGCTGCTGCTGACCCGCGACGGCGGCAAGGA 1161 1200
Leu122-Ser199	(1131)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Val127-Asn195	(1161)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ile201B	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ala204	(1113)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Ile201	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Val120-Thr202	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Lys121-Val200	(1125)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC
Consensus	(1161)	GATCAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGC 1201 1240
Leu122-Ser199	(1171)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val127-Asn195	(1201)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201B	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ala204	(1153)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ile201	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Thr202	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Lys121-Val200	(1165)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Consensus	(1201)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA 1241 1280
Leu122-Ser199	(1211)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCACAA
Val127-Asn195	(1241)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCACAA
Val120-Ile201B	(1199)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCACAA
Val120-Ala204	(1193)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCACAA
Val120-Ile201	(1199)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCACAA

FIG. 3E

Val120-Thr202	(1199)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAA
Lys121-Val200	(1205)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAA
Consensus	(1241)	AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAA 1281 1320
Leu122-Ser199	(1251)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Val127-Asn195	(1281)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Val120-Ile201B	(1239)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Val120-Ala204	(1233)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Val120-Ile201	(1239)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Val120-Thr202	(1239)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Lys121-Val200	(1245)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG
Consensus	(1281)	GGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTG 1321 1360
Leu122-Ser199	(1291)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Val127-Asn195	(1321)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Val120-Ile201B	(1279)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Val120-Ala204	(1273)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Val120-Ile201	(1279)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Val120-Thr202	(1279)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Lys121-Val200	(1285)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG
Consensus	(1321)	ACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCG 1361 1400
Leu122-Ser199	(1331)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Val127-Asn195	(1361)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Val120-Ile201B	(1319)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Val120-Ala204	(1313)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Val120-Ile201	(1319)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Val120-Thr202	(1319)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Lys121-Val200	(1325)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
Consensus	(1361)	GCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA 1401 1440
Leu122-Ser199	(1371)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Val127-Asn195	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Val120-Ile201B	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Val120-Ala204	(1353)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Val120-Ile201	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Val120-Thr202	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Lys121-Val200	(1365)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC
Consensus	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAAC 1441 1480
Leu122-Ser199	(1411)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Val127-Asn195	(1441)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Val120-Ile201B	(1399)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Val120-Ala204	(1393)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Val120-Ile201	(1399)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Val120-Thr202	(1399)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Lys121-Val200	(1405)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC
Consensus	(1441)	AACCTGCTGCGCGCCATCGAGGGCCAGCAGCACCTGCTGC 1481 1520
Leu122-Ser199	(1451)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Val127-Asn195	(1481)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Val120-Ile201B	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Val120-Ala204	(1433)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Val120-Ile201	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Val120-Thr202	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Lys121-Val200	(1445)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT
Consensus	(1481)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT

FIG. 3F

		1521		1560
Leu122-Ser199	(1491)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Val127-Asn195	(1521)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Val120-Ile201B	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Val120-Ala204	(1473)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Val120-Ile201	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Val120-Thr202	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Lys121-Val200	(1485)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
Consensus	(1521)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG		
		1561		1600
Leu122-Ser199	(1531)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Val127-Asn195	(1561)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Val120-Ile201B	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Val120-Ala204	(1513)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Val120-Ile201	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Val120-Thr202	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Lys121-Val200	(1525)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
Consensus	(1561)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG		
		1601		1640
Leu122-Ser199	(1571)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Val127-Asn195	(1601)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Val120-Ile201B	(1559)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Val120-Ala204	(1553)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Val120-Ile201	(1559)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Val120-Thr202	(1559)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Lys121-Val200	(1565)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
Consensus	(1601)	CCGTGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGA		
		1641		1680
Leu122-Ser199	(1611)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Val127-Asn195	(1641)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Val120-Ile201B	(1599)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Val120-Ala204	(1593)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Val120-Ile201	(1599)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Val120-Thr202	(1599)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Lys121-Val200	(1605)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
Consensus	(1641)	CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC		
		1681		1720
Leu122-Ser199	(1651)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Val127-Asn195	(1681)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Val120-Ile201B	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Val120-Ala204	(1633)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Val120-Ile201	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Val120-Thr202	(1639)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Lys121-Val200	(1645)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
Consensus	(1681)	GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCG		
		1721		1760
Leu122-Ser199	(1691)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Val127-Asn195	(1721)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Val120-Ile201B	(1679)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Val120-Ala204	(1673)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Val120-Ile201	(1679)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Val120-Thr202	(1679)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Lys121-Val200	(1685)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
Consensus	(1721)	AGGAGAGCCAGAACCCAGCAGGAGAGGAACGAGCAGGAGCT		
		1761		1800
Leu122-Ser199	(1731)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAAGTGGTTC		
Val127-Asn195	(1761)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAAGTGGTTC		

FIG. 3G

Val120-Ile201B	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC
Val120-Ala204	(1713)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC
Val120-Ile201	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC
Val120-Thr202	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC
Lys121-Val200	(1725)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC
Consensus	(1761)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTC 1801 1840
Leu122-Ser199	(1771)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val127-Asn195	(1801)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ile201B	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ala204	(1753)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Ile201	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Val120-Thr202	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Lys121-Val200	(1765)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA
Consensus	(1801)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCA 1841 1880
Leu122-Ser199	(1811)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Val127-Asn195	(1841)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Val120-Ile201B	(1799)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Val120-Ala204	(1793)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Val120-Ile201	(1799)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Val120-Thr202	(1799)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Lys121-Val200	(1805)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC
Consensus	(1841)	TGATCGTGGGCGGCCCTGGTGGGCTGCGCATCGTGTTCAC 1881 1920
Leu122-Ser199	(1851)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Val127-Asn195	(1881)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Val120-Ile201B	(1839)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Val120-Ala204	(1833)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Val120-Ile201	(1839)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Val120-Thr202	(1839)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Lys121-Val200	(1845)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC
Consensus	(1881)	CGTGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGC 1921 1960
Leu122-Ser199	(1891)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Val127-Asn195	(1921)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Val120-Ile201B	(1879)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Val120-Ala204	(1873)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Val120-Ile201	(1879)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Val120-Thr202	(1879)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Lys121-Val200	(1885)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC
Consensus	(1921)	CCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGC 1961 2000
Leu122-Ser199	(1931)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Val127-Asn195	(1961)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Val120-Ile201B	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Val120-Ala204	(1913)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Val120-Ile201	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Val120-Thr202	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Lys121-Val200	(1925)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG
Consensus	(1961)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGGGGAGCG 2001 2040
Leu122-Ser199	(1971)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val127-Asn195	(2001)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ile201B	(1959)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ala204	(1953)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Val120-Ile201	(1959)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG

FIG. 3H

Vall120-Thr202	(1959)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Lys121-Val200	(1965)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
Consensus	(2001)	CGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG
		2041 2080
Leu122-Ser199	(2011)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall127-Asn195	(2041)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall120-Ile201B	(1999)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall120-Ala204	(1993)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall120-Ile201	(1999)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall120-Thr202	(1999)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Lys121-Val200	(2005)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Consensus	(2041)	GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
		2081 2120
Leu122-Ser199	(2051)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Vall127-Asn195	(2081)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Vall120-Ile201B	(2039)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Vall120-Ala204	(2033)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Vall120-Ile201	(2039)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Vall120-Thr202	(2039)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Lys121-Val200	(2045)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
Consensus	(2081)	GCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
		2121 2160
Leu122-Ser199	(2091)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Vall127-Asn195	(2121)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Vall120-Ile201B	(2079)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Vall120-Ala204	(2073)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Vall120-Ile201	(2079)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Vall120-Thr202	(2079)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Lys121-Val200	(2085)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
Consensus	(2121)	CATCGTGGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTG
		2161 2200
Leu122-Ser199	(2131)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall127-Asn195	(2161)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall120-Ile201B	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall120-Ala204	(2113)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall120-Ile201	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall120-Thr202	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Lys121-Val200	(2125)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Consensus	(2161)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
		2201 2240
Leu122-Ser199	(2171)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall127-Asn195	(2201)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall120-Ile201B	(2159)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall120-Ala204	(2153)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall120-Ile201	(2159)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall120-Thr202	(2159)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Lys121-Val200	(2165)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
Consensus	(2201)	TGAAGAACAGGGCGCTGAGCCTGTTGACGCCATCGCCAT
		2241 2280
Leu122-Ser199	(2211)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Vall127-Asn195	(2241)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Vall120-Ile201B	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Vall120-Ala204	(2193)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Vall120-Ile201	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Vall120-Thr202	(2199)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Lys121-Val200	(2205)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
Consensus	(2241)	CGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC

FIG. 3I

		2281		2320
Leu122-Ser199	(2251)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Val127-Asn195	(2281)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ile201B	(2239)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ala204	(2233)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ile201	(2239)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Thr202	(2239)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Lys121-Val200	(2245)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
Consensus	(2281)	CAGCGCATCGGCCGGCGCCTTCCTGCACATCCCCCGCCGCA		
		2321		2360
Leu122-Ser199	(2291)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAGCG		
Val127-Asn195	(2321)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAG--		
Val120-Ile201B	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAGCG		
Val120-Ala204	(2273)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAG--		
Val120-Ile201	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAG--		
Val120-Thr202	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAG--		
Lys121-Val200	(2285)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAGCG		
Consensus	(2321)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAATCGAG		
		2361		
Leu122-Ser199	(2331)	TGCT		
Val127-Asn195	(2359)	----		
Val120-Ile201B	(2319)	TGCT		
Val120-Ala204	(2311)	----		
Val120-Ile201	(2317)	----		
Val120-Thr202	(2317)	----		
Lys121-Val200	(2325)	TGCT		
Consensus	(2361)			

FIG. 3J

	1	40
Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCT
Trp427-Gly431	(1)	
Gln422-Tyr435B	(1)	
Arg426-Gly431	(1)	
Ile423-Met434	(1)	
Gln422-Tyr435	(1)	
Arg426-Lys432	(1)	
Arg426-Gly431B	(1)	
Asn425-Lys432	(1)	
Consensus	(1)	
	41	80
Ile424-Ala433	(41)	
Trp427-Gly431	(41)	
Gln422-Tyr435B	(41)	
Arg426-Gly431	(41)	
Ile423-Met434	(41)	
Gln422-Tyr435	(41)	
Arg426-Lys432	(41)	
Arg426-Gly431B	(41)	
Asn425-Lys432	(41)	
Consensus	(41)	
	81	120
Ile424-Ala433	(81)	
Trp427-Gly431	(81)	
Gln422-Tyr435B	(81)	
Arg426-Gly431	(81)	
Ile423-Met434	(81)	
Gln422-Tyr435	(81)	
Arg426-Lys432	(81)	
Arg426-Gly431B	(81)	
Asn425-Lys432	(81)	
Consensus	(81)	
	121	160
Ile424-Ala433	(121)	
Trp427-Gly431	(121)	
Gln422-Tyr435B	(121)	
Arg426-Gly431	(121)	
Ile423-Met434	(121)	
Gln422-Tyr435	(121)	
Arg426-Lys432	(121)	
Arg426-Gly431B	(121)	
Asn425-Lys432	(121)	
Consensus	(121)	
	161	200
Ile424-Ala433	(161)	
Trp427-Gly431	(161)	
Gln422-Tyr435B	(161)	
Arg426-Gly431	(161)	
Ile423-Met434	(161)	
Gln422-Tyr435	(161)	
Arg426-Lys432	(161)	
Arg426-Gly431B	(161)	
Asn425-Lys432	(161)	
Consensus	(161)	
	201	240
Ile424-Ala433	(201)	

FIG. 4A

FIG. 4B

Arg426-Gly431	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(401)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(401)	ACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGATGGA
		441 480
Ile424-Ala433	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(441)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(441)	CCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACC
		481 520
Ile424-Ala433	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(481)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(481)	AGCATCCGCAACAAGATGCAGAAGGAGTACGCCCTGTTCT
		521 560
Ile424-Ala433	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(521)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(521)	ACAAGCTGGACGTGGTGGCCATCGACAACGACAACACCAG
		561 600
Ile424-Ala433	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Lys432	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431B	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Asn425-Lys432	(561)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Consensus	(561)	CTACAAGCTGATCAACTGCAACACCAGCGTGATCACCAG
		601 640
Ile424-Ala433	(601)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Trp427-Gly431	(601)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Gln422-Tyr435B	(601)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Arg426-Gly431	(601)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA
Ile423-Met434	(601)	AGCGGACGAAACAGCAAGAGCAGCAACTGGAAGGAGATGGA

FIG. 4C

Gln422-Tyr435	(601)	GCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACT
Arg426-Lys432	(601)	641
Arg426-Gly431B	(601)	680
Asn425-Lys432	(601)	
Consensus	(601)	
Ile424-Ala433	(641)	
Trp427-Gly431	(641)	
Gln422-Tyr435B	(641)	
Arg426-Gly431	(641)	
Ile423-Met434	(641)	
Gln422-Tyr435	(641)	
Arg426-Lys432	(641)	
Arg426-Gly431B	(641)	
Asn425-Lys432	(641)	
Consensus	(641)	ACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGA
Ile424-Ala433	(681)	681
Trp427-Gly431	(681)	720
Gln422-Tyr435B	(681)	
Arg426-Gly431	(681)	
Ile423-Met434	(681)	
Gln422-Tyr435	(681)	
Arg426-Lys432	(681)	
Arg426-Gly431B	(681)	
Asn425-Lys432	(681)	
Consensus	(681)	CAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGC
Ile424-Ala433	(721)	721
Trp427-Gly431	(721)	760
Gln422-Tyr435B	(721)	
Arg426-Gly431	(721)	
Ile423-Met434	(721)	
Gln422-Tyr435	(721)	
Arg426-Lys432	(721)	
Arg426-Gly431B	(721)	
Asn425-Lys432	(721)	
Consensus	(721)	ACCGTGCAAGTGACCCACGGCATCCGCCCGTGGTGAGCA
Ile424-Ala433	(761)	761
Trp427-Gly431	(761)	800
Gln422-Tyr435B	(761)	
Arg426-Gly431	(761)	
Ile423-Met434	(761)	
Gln422-Tyr435	(761)	
Arg426-Lys432	(761)	
Arg426-Gly431B	(761)	
Asn425-Lys432	(761)	
Consensus	(761)	CCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGT
Ile424-Ala433	(801)	801
Trp427-Gly431	(801)	840
Gln422-Tyr435B	(801)	
Arg426-Gly431	(801)	
Ile423-Met434	(801)	
Gln422-Tyr435	(801)	
Arg426-Lys432	(801)	

FIG. 4D

Arg426-Gly431B	(801)	GGTATCCGCGAGCGAGA	841	880
Asn425-Lys432	(801)	GGTATCCGCGAGCGAGA		
Consensus	(801)	GGTATCCGCGAGCGAGA		
Ile424-Ala433	(841)	ATCATCGTGCAGCTGA	881	920
Trp427-Gly431	(841)	ATCATCGTGCAGCTGA		
Gln422-Tyr435B	(841)	ATCATCGTGCAGCTGA		
Arg426-Gly431	(841)	ATCATCGTGCAGCTGA		
Ile423-Met434	(841)	ATCATCGTGCAGCTGA		
Gln422-Tyr435	(841)	ATCATCGTGCAGCTGA		
Arg426-Lys432	(841)	ATCATCGTGCAGCTGA		
Arg426-Gly431B	(841)	ATCATCGTGCAGCTGA		
Asn425-Lys432	(841)	ATCATCGTGCAGCTGA		
Consensus	(841)	ATCATCGTGCAGCTGA		
Ile424-Ala433	(881)	ATCATCGTGCAGCTGA	881	920
Trp427-Gly431	(881)	ATCATCGTGCAGCTGA		
Gln422-Tyr435B	(881)	ATCATCGTGCAGCTGA		
Arg426-Gly431	(881)	ATCATCGTGCAGCTGA		
Ile423-Met434	(881)	ATCATCGTGCAGCTGA		
Gln422-Tyr435	(881)	ATCATCGTGCAGCTGA		
Arg426-Lys432	(881)	ATCATCGTGCAGCTGA		
Arg426-Gly431B	(881)	ATCATCGTGCAGCTGA		
Asn425-Lys432	(881)	ATCATCGTGCAGCTGA		
Consensus	(881)	ATCATCGTGCAGCTGA		
Ile424-Ala433	(921)	ATCATCGTGCAGCTGA	921	960
Trp427-Gly431	(921)	ATCATCGTGCAGCTGA		
Gln422-Tyr435B	(921)	ATCATCGTGCAGCTGA		
Arg426-Gly431	(921)	ATCATCGTGCAGCTGA		
Ile423-Met434	(921)	ATCATCGTGCAGCTGA		
Gln422-Tyr435	(921)	ATCATCGTGCAGCTGA		
Arg426-Lys432	(921)	ATCATCGTGCAGCTGA		
Arg426-Gly431B	(921)	ATCATCGTGCAGCTGA		
Asn425-Lys432	(921)	ATCATCGTGCAGCTGA		
Consensus	(921)	ATCATCGTGCAGCTGA		
Ile424-Ala433	(961)	ATCATCGTGCAGCTGA	961	1000
Trp427-Gly431	(961)	ATCATCGTGCAGCTGA		
Gln422-Tyr435B	(961)	ATCATCGTGCAGCTGA		
Arg426-Gly431	(961)	ATCATCGTGCAGCTGA		
Ile423-Met434	(961)	ATCATCGTGCAGCTGA		
Gln422-Tyr435	(961)	ATCATCGTGCAGCTGA		
Arg426-Lys432	(961)	ATCATCGTGCAGCTGA		
Arg426-Gly431B	(961)	ATCATCGTGCAGCTGA		
Asn425-Lys432	(961)	ATCATCGTGCAGCTGA		
Consensus	(961)	ATCATCGTGCAGCTGA		
Ile424-Ala433	(1001)	ATCATCGTGCAGCTGA	1001	1040
Trp427-Gly431	(1001)	ATCATCGTGCAGCTGA		
Gln422-Tyr435B	(1001)	ATCATCGTGCAGCTGA		
Arg426-Gly431	(1001)	ATCATCGTGCAGCTGA		
Ile423-Met434	(1001)	ATCATCGTGCAGCTGA		
Gln422-Tyr435	(1001)	ATCATCGTGCAGCTGA		
Arg426-Lys432	(1001)	ATCATCGTGCAGCTGA		
Arg426-Gly431B	(1001)	ATCATCGTGCAGCTGA		
Asn425-Lys432	(1001)	ATCATCGTGCAGCTGA		

FIG. 4E

Consensus	(1001)	GGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGC	1041	1080
Ile424-Ala433	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Trp427-Gly431	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435B	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Ile423-Met434	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Lys432	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431B	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Asn425-Lys432	(1041)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Consensus	(1041)	CCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGC	1081	1120
Ile424-Ala433	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Trp427-Gly431	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435B	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Ile423-Met434	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Lys432	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431B	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Asn425-Lys432	(1081)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Consensus	(1081)	GGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG	1121	1160
Ile424-Ala433	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Trp427-Gly431	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435B	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Ile423-Met434	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Lys432	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431B	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Asn425-Lys432	(1121)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Consensus	(1121)	GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAA	1161	1200
Ile424-Ala433	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Trp427-Gly431	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435B	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Ile423-Met434	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Lys432	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431B	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Asn425-Lys432	(1161)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Consensus	(1161)	CAGCACCTGGAACAACACCATCGGCCCAACAACACCAAC	1201	1240
Ile424-Ala433	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Trp427-Gly431	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435B	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Ile423-Met434	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Gln422-Tyr435	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Lys432	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Arg426-Gly431B	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Asn425-Lys432	(1201)	GCACCTGCGGCTGATGAGTTCCTGCTGAGGAGAGGAGG		
Consensus	(1201)	GGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCA	1241	1280

FIG. 4F

FIG. 4G

Gln422-Tyr435B	(1417)	AGCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGG
Arg426-Gly431	(1441)	1481
Ile423-Met434	(1423)	1520
Gln422-Tyr435	(1417)	
Arg426-Lys432	(1441)	
Arg426-Gly431B	(1441)	
Asn425-Lys432	(1435)	
Consensus	(1441)	CCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGG
Ile424-Ala433	(1469)	
Trp427-Gly431	(1481)	
Gln422-Tyr435B	(1457)	
Arg426-Gly431	(1481)	
Ile423-Met434	(1463)	
Gln422-Tyr435	(1457)	
Arg426-Lys432	(1481)	
Arg426-Gly431B	(1481)	
Asn425-Lys432	(1475)	
Consensus	(1481)	TGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTT
Ile424-Ala433	(1509)	1521
Trp427-Gly431	(1521)	1560
Gln422-Tyr435B	(1497)	
Arg426-Gly431	(1521)	
Ile423-Met434	(1503)	
Gln422-Tyr435	(1497)	
Arg426-Lys432	(1521)	
Arg426-Gly431B	(1521)	
Asn425-Lys432	(1515)	
Consensus	(1521)	CCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCC
Ile424-Ala433	(1549)	1561
Trp427-Gly431	(1561)	1600
Gln422-Tyr435B	(1537)	
Arg426-Gly431	(1561)	
Ile423-Met434	(1543)	
Gln422-Tyr435	(1537)	
Arg426-Lys432	(1561)	
Arg426-Gly431B	(1561)	
Asn425-Lys432	(1555)	
Consensus	(1561)	CGCAGCCTGACCCTGACCGTGCAGGCCCGCCAGCTGCTGA
Ile424-Ala433	(1589)	1601
Trp427-Gly431	(1601)	1640
Gln422-Tyr435B	(1577)	
Arg426-Gly431	(1601)	
Ile423-Met434	(1583)	
Gln422-Tyr435	(1577)	
Arg426-Lys432	(1601)	
Arg426-Gly431B	(1601)	
Asn425-Lys432	(1595)	
Consensus	(1601)	GCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGGCCAT
Ile424-Ala433	(1629)	1641
Trp427-Gly431	(1641)	1680
Gln422-Tyr435B	(1617)	
Arg426-Gly431	(1641)	

FIG. 4H

FIG. 4I

Arg426-Lys432	(1841)	TCACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACAC
Arg426-Gly431B	(1841)	TCACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACAC
Asn425-Lys432	(1835)	TCACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACAC
Consensus	(1841)	TCACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACAC
		1881 1920
Ile424-Ala433	(1869)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Trp427-Gly431	(1881)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Gln422-Tyr435B	(1857)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Arg426-Gly431	(1881)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Ile423-Met434	(1863)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Gln422-Tyr435	(1857)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Arg426-Lys432	(1881)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Arg426-Gly431B	(1881)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Asn425-Lys432	(1875)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
Consensus	(1881)	CAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAG
		1921 1960
Ile424-Ala433	(1909)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Trp427-Gly431	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Gln422-Tyr435B	(1897)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Gly431	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Ile423-Met434	(1903)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Gln422-Tyr435	(1897)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Lys432	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Arg426-Gly431B	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Asn425-Lys432	(1915)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
Consensus	(1921)	CAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGT
		1961 2000
Ile424-Ala433	(1949)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Trp427-Gly431	(1961)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Gln422-Tyr435B	(1937)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Arg426-Gly431	(1961)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Ile423-Met434	(1943)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Gln422-Tyr435	(1937)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Arg426-Lys432	(1961)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Arg426-Gly431B	(1961)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Asn425-Lys432	(1955)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
Consensus	(1961)	GGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCT
		2001 2040
Ile424-Ala433	(1989)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Trp427-Gly431	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Gln422-Tyr435B	(1977)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Gly431	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Ile423-Met434	(1983)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Gln422-Tyr435	(1977)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Lys432	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Gly431B	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Asn425-Lys432	(1995)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Consensus	(2001)	GTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
		2041 2080
Ile424-Ala433	(2029)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Trp427-Gly431	(2041)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Gln422-Tyr435B	(2017)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Gly431	(2041)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Ile423-Met434	(2023)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Gln422-Tyr435	(2017)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Lys432	(2041)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
Arg426-Gly431B	(2041)	CTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG

FIG. 4J

Asn425-Lys432	(2035)	GTGGGCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGA
Consensus	(2041)	2081 2120
Ile424-Ala433	(2069)	
Trp427-Gly431	(2081)	
Gln422-Tyr435B	(2057)	
Arg426-Gly431	(2081)	
Ile423-Met434	(2063)	
Gln422-Tyr435	(2057)	
Arg426-Lys432	(2081)	
Arg426-Gly431B	(2081)	
Asn425-Lys432	(2075)	
Consensus	(2081)	ACC GCGT GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGAC
		2121 2160
Ile424-Ala433	(2109)	
Trp427-Gly431	(2121)	
Gln422-Tyr435B	(2097)	
Arg426-Gly431	(2121)	
Ile423-Met434	(2103)	
Gln422-Tyr435	(2097)	
Arg426-Lys432	(2121)	
Arg426-Gly431B	(2121)	
Asn425-Lys432	(2115)	
Consensus	(2121)	CCGCTTCCCCGCCCCCGCGCCCCGACCGCCCCGAGGGC
		2161 2200
Ile424-Ala433	(2149)	
Trp427-Gly431	(2161)	
Gln422-Tyr435B	(2137)	
Arg426-Gly431	(2161)	
Ile423-Met434	(2143)	
Gln422-Tyr435	(2137)	
Arg426-Lys432	(2161)	
Arg426-Gly431B	(2161)	
Asn425-Lys432	(2155)	
Consensus	(2161)	ATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCA
		2201 2240
Ile424-Ala433	(2189)	
Trp427-Gly431	(2201)	
Gln422-Tyr435B	(2177)	
Arg426-Gly431	(2201)	
Ile423-Met434	(2183)	
Gln422-Tyr435	(2177)	
Arg426-Lys432	(2201)	
Arg426-Gly431B	(2201)	
Asn425-Lys432	(2195)	
Consensus	(2201)	GCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA
		2241 2280
Ile424-Ala433	(2229)	
Trp427-Gly431	(2241)	
Gln422-Tyr435B	(2217)	
Arg426-Gly431	(2241)	
Ile423-Met434	(2223)	
Gln422-Tyr435	(2217)	
Arg426-Lys432	(2241)	
Arg426-Gly431B	(2241)	
Asn425-Lys432	(2235)	
Consensus	(2241)	CCTGCGCAGCCTGTGCCTGTTTCAGCTACCACCGCCTGCGC

FIG. 4K

		2281	2320
Ile424-Ala433	(2269)	GACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGG	
Trp427-Gly431	(2281)		
Gln422-Tyr435B	(2257)		
Arg426-Gly431	(2281)		
Ile423-Met434	(2263)		
Gln422-Tyr435	(2257)		
Arg426-Lys432	(2281)		
Arg426-Gly431B	(2281)		
Asn425-Lys432	(2275)		
Consensus	(2281)	GACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGG	
		2321	2360
Ile424-Ala433	(2309)		
Trp427-Gly431	(2321)		
Gln422-Tyr435B	(2297)		
Arg426-Gly431	(2321)		
Ile423-Met434	(2303)		
Gln422-Tyr435	(2297)		
Arg426-Lys432	(2321)		
Arg426-Gly431B	(2321)		
Asn425-Lys432	(2315)		
Consensus	(2321)	GCCGCCGCCGCTGGGAGGCCCTGAAGTACTGGGGCAACCT	
		2361	2400
Ile424-Ala433	(2349)		
Trp427-Gly431	(2361)		
Gln422-Tyr435B	(2337)		
Arg426-Gly431	(2361)		
Ile423-Met434	(2343)		
Gln422-Tyr435	(2337)		
Arg426-Lys432	(2361)		
Arg426-Gly431B	(2361)		
Asn425-Lys432	(2355)		
Consensus	(2361)	GCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG	
		2401	2440
Ile424-Ala433	(2389)		
Trp427-Gly431	(2401)		
Gln422-Tyr435B	(2377)		
Arg426-Gly431	(2401)		
Ile423-Met434	(2383)		
Gln422-Tyr435	(2377)		
Arg426-Lys432	(2401)		
Arg426-Gly431B	(2401)		
Asn425-Lys432	(2395)		
Consensus	(2401)	AGCCTGTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA	
		2441	2480
Ile424-Ala433	(2429)		
Trp427-Gly431	(2441)		
Gln422-Tyr435B	(2417)		
Arg426-Gly431	(2441)		
Ile423-Met434	(2423)		
Gln422-Tyr435	(2417)		
Arg426-Lys432	(2441)		
Arg426-Gly431B	(2441)		
Asn425-Lys432	(2435)		
Consensus	(2441)	CCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGC	
		2481	2520
Ile424-Ala433	(2469)		

FIG. 4L

FIG. 4M

		30
Leu122-Ser199-Tryp427-Gly431	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Vall127-Asn195-Arg426-Gly431	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Vall120-Thr202-Ile424-Ala433	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Lys432	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Gly431	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Lys121-Val200-Asn425-Lys432	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Vall120-Ile201-Ile424-Ala433	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Vall120-Ile201B-Ile424-Ala433	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
Consensus	(1)	GAATTGCGCCACCATGGATGCAATGAAGAGA
	31	60
Leu122-Ser199-Tryp427-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Vall127-Asn195-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Vall120-Thr202-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Lys121-Val200-Asn425-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Vall120-Ile201-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Vall120-Ile201B-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Consensus	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
	61	90
Leu122-Ser199-Tryp427-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Vall127-Asn195-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Vall120-Thr202-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Leu122-Ser199-Arg426-Lys432	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Leu122-Ser199-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Lys121-Val200-Asn425-Lys432	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Vall120-Ile201-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Vall120-Ile201B-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Consensus	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
	91	120
Leu122-Ser199-Tryp427-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Vall127-Asn195-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Vall120-Thr202-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Lys432	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Leu122-Ser199-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Lys121-Val200-Asn425-Lys432	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Vall120-Ile201-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Vall120-Ile201B-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
Consensus	(91)	AAGCTGTGGGTGACCGTGTACTACGGCGTG
	121	150
Leu122-Ser199-Tryp427-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Vall127-Asn195-Arg426-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Vall120-Thr202-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Leu122-Ser199-Arg426-Lys432	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Leu122-Ser199-Arg426-Gly431	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Lys121-Val200-Asn425-Lys432	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Vall120-Ile201-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Vall120-Ile201B-Ile424-Ala433	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
Consensus	(121)	CCCGTGTGGAAGGAGGCCACCACCACCGTG
	151	180
Leu122-Ser199-Tryp427-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Vall127-Asn195-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Vall120-Thr202-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Lys121-Val200-Asn425-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC

FIG. 5A

WO 00/39303	29	/	65	PCT/US99/31272
Vall20-Ile201-Ile424-Ala433	(151)			TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Vall20-Ile201B-Ile424-Ala433	(151)			TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Consensus	(151)			TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leul22-Ser199-Tryp427-Gly431	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Vall27-Asn195-Arg426-Gly431	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Vall20-Thr202-Ile424-Ala433	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Leul22-Ser199-Arg426-Lys432	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Leul22-Ser199-Arg426-Gly431	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Lys121-Val200-Asn425-Lys432	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Vall20-Ile201-Ile424-Ala433	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Vall20-Ile201B-Ile424-Ala433	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Consensus	(181)			ACCGAGGTGCACAACGTGTGGGCCACCCAC
Leul22-Ser199-Tryp427-Gly431	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Vall27-Asn195-Arg426-Gly431	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Vall20-Thr202-Ile424-Ala433	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Leul22-Ser199-Arg426-Lys432	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Leul22-Ser199-Arg426-Gly431	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Lys121-Val200-Asn425-Lys432	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Vall20-Ile201-Ile424-Ala433	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Vall20-Ile201B-Ile424-Ala433	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Consensus	(211)			GCCTGCGTGCCCAACCGACCCCAACCCCCAG
Leul22-Ser199-Tryp427-Gly431	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Vall27-Asn195-Arg426-Gly431	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Vall20-Thr202-Ile424-Ala433	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Leul22-Ser199-Arg426-Lys432	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Leul22-Ser199-Arg426-Gly431	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Lys121-Val200-Asn425-Lys432	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Vall20-Ile201-Ile424-Ala433	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Vall20-Ile201B-Ile424-Ala433	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Consensus	(241)			GAGATCGTGCTGGAGAACGTGACCGAGAAC
Leul22-Ser199-Tryp427-Gly431	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Vall27-Asn195-Arg426-Gly431	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Vall20-Thr202-Ile424-Ala433	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Leul22-Ser199-Arg426-Lys432	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Leul22-Ser199-Arg426-Gly431	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Lys121-Val200-Asn425-Lys432	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Vall20-Ile201-Ile424-Ala433	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Vall20-Ile201B-Ile424-Ala433	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Consensus	(271)			TTCAACATGTGGAGAACAACATGGTGGAG
Leul22-Ser199-Tryp427-Gly431	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Vall27-Asn195-Arg426-Gly431	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Vall20-Thr202-Ile424-Ala433	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Leul22-Ser199-Arg426-Lys432	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Leul22-Ser199-Arg426-Gly431	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Lys121-Val200-Asn425-Lys432	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Vall20-Ile201-Ile424-Ala433	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Vall20-Ile201B-Ile424-Ala433	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Consensus	(301)			CAGATGCACGAGGACATCATCAGCCTGTGG
Leul22-Ser199-Tryp427-Gly431	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Vall27-Asn195-Arg426-Gly431	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Vall20-Thr202-Ile424-Ala433	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG

FIG. 5B

WO 00/39303	30	/	65	PCT/US99/31272
Leu122-Ser199-Arg426-Lys432	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Arg426-Gly431	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Lys121-Val200-Asn425-Lys432	(331)			GACCAGAGCCTGAAGCCCTGCGTGAA----
Val120-Ile201-Ile424-Ala433	(331)			GACCAGAGCCTGAAGCCCTGCGTG-----
Val120-Ile201B-Ile424-Ala433	(331)			GACCAGAGCCTGAAGCCCTGCGTG-----
Consensus	(331)			GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
				361 390
Leu122-Ser199-Tryp427-Gly431	(361)			-----GG-----
Val127-Asn195-Arg426-Gly431	(361)			ACCCCCCTGTGCGTGGGGCAGGGAAGTGC
Val120-Thr202-Ile424-Ala433	(355)			-----GG-----
Leu122-Ser199-Arg426-Lys432	(361)			-----GG-----
Leu122-Ser199-Arg426-Gly431	(361)			-----GG-----
Lys121-Val200-Asn425-Lys432	(357)			-----GG-----
Val120-Ile201-Ile424-Ala433	(355)			-----
Val120-Ile201B-Ile424-Ala433	(355)			-----
Consensus	(361)			GG
				391 420
Leu122-Ser199-Tryp427-Gly431	(363)			--CAACAGCGTGATCACCCAGGCCTGCCCC
Val127-Asn195-Arg426-Gly431	(391)			AACACAGCGTGATCACCCAGGCCTGCCCC
Val120-Thr202-Ile424-Ala433	(357)			-----CGGCGC---CACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Lys432	(363)			--CAACAGCGTGATCACCCAGGCCTGCCCC
Leu122-Ser199-Arg426-Gly431	(363)			--CAACAGCGTGATCACCCAGGCCTGCCCC
Lys121-Val200-Asn425-Lys432	(359)			----CECCGGTGATCACCCAGGCCTGCCCC
Val120-Ile201-Ile424-Ala433	(355)			-----GCGGCGATCACCCAGGCCTGCCCC
Val120-Ile201B-Ile424-Ala433	(355)			-----CCGCGCATCACCCAGGCCTGCCCC
Consensus	(391)			CA CAGCGTGATCACCCAGGCCTGCCCC
				421 450
Leu122-Ser199-Tryp427-Gly431	(391)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val127-Asn195-Arg426-Gly431	(421)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Thr202-Ile424-Ala433	(379)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Leu122-Ser199-Arg426-Lys432	(391)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Leu122-Ser199-Arg426-Gly431	(391)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Lys121-Val200-Asn425-Lys432	(385)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Ile201-Ile424-Ala433	(379)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Val120-Ile201B-Ile424-Ala433	(379)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
Consensus	(421)			AAGGTGAGCTTCGAGCCCATCCCCATCCAC
				451 480
Leu122-Ser199-Tryp427-Gly431	(421)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val127-Asn195-Arg426-Gly431	(451)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Thr202-Ile424-Ala433	(409)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Lys432	(421)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Gly431	(421)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Lys121-Val200-Asn425-Lys432	(415)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Ile201-Ile424-Ala433	(409)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Ile201B-Ile424-Ala433	(409)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
Consensus	(451)			TACTGCGCCCGCGCGGCTTCGCCATCCTG
				481 510
Leu122-Ser199-Tryp427-Gly431	(451)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val127-Asn195-Arg426-Gly431	(481)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-Thr202-Ile424-Ala433	(439)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Lys432	(451)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Leu122-Ser199-Arg426-Gly431	(451)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Lys121-Val200-Asn425-Lys432	(445)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-Ile201-Ile424-Ala433	(439)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Val120-Ile201B-Ile424-Ala433	(439)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
Consensus	(481)			AAGTGCAACGACAAGAAGTTCAACGGCAGC
				511 540

FIG. 5C

Leu122-Ser199-Tryp427-Gly431	(481)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Val127-Asn195-Arg426-Gly431	(511)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Val120-Thr202-Ile424-Ala433	(469)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Leu122-Ser199-Arg426-Lys432	(481)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Leu122-Ser199-Arg426-Gly431	(481)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Lys121-Val200-Asn425-Lys432	(475)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Val120-Ile201-Ile424-Ala433	(469)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Val120-Ile201B-Ile424-Ala433	(469)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Consensus	(511)	GGCCCCGTCACCAACGTGAGCACCCTGCGAG
Leu122-Ser199-Tryp427-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val127-Asn195-Arg426-Gly431	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Thr202-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Arg426-Lys432	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Ile201-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Val120-Ile201B-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Consensus	(541)	TGCACCCACGGCATCCGCCCCGTGGTGAGC
Leu122-Ser199-Tryp427-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val127-Asn195-Arg426-Gly431	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Thr202-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200-Asn425-Lys432	(535)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val120-Ile201B-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Consensus	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Tryp427-Gly431	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val120-Thr202-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Arg426-Lys432	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Arg426-Gly431	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Lys121-Val200-Asn425-Lys432	(565)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val120-Ile201-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val120-Ile201B-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Consensus	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Tryp427-Gly431	(601)	TTCACCGGACACGCCAAGACCATCATCGTG
Val127-Asn195-Arg426-Gly431	(631)	TTCACCGGACACGCCAAGACCATCATCGTG
Val120-Thr202-Ile424-Ala433	(589)	TTCACCGGACACGCCAAGACCATCATCGTG
Leu122-Ser199-Arg426-Lys432	(601)	TTCACCGGACACGCCAAGACCATCATCGTG
Leu122-Ser199-Arg426-Gly431	(601)	TTCACCGGACACGCCAAGACCATCATCGTG
Lys121-Val200-Asn425-Lys432	(595)	TTCACCGGACACGCCAAGACCATCATCGTG
Val120-Ile201-Ile424-Ala433	(589)	TTCACCGGACACGCCAAGACCATCATCGTG
Val120-Ile201B-Ile424-Ala433	(589)	TTCACCGGACACGCCAAGACCATCATCGTG
Consensus	(631)	TTCACCGGACACGCCAAGACCATCATCGTG
Leu122-Ser199-Tryp427-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val127-Asn195-Arg426-Gly431	(661)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val120-Thr202-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Lys432	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Lys121-Val200-Asn425-Lys432	(625)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val120-Ile201-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC

FIG. 5D

Val120-Ile201B-Ile424-Ala433	(619) CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Consensus	(661) CAGCTGAAGGAGAGCGTGGAGATCAACTGC
	691 720
Leu122-Ser199-Tryp427-Gly431	(661) ACCCGCCCCAACAACAACACCCGCAAGAGC
Val127-Asn195-Arg426-Gly431	(691) ACCCGCCCCAACAACAACACCCGCAAGAGC
Val120-Thr202-Ile424-Ala433	(649) ACCCGCCCCAACAACAACACCCGCAAGAGC
Leu122-Ser199-Arg426-Lys432	(661) ACCCGCCCCAACAACAACACCCGCAAGAGC
Leu122-Ser199-Arg426-Gly431	(661) ACCCGCCCCAACAACAACACCCGCAAGAGC
Lys121-Val200-Asn425-Lys432	(655) ACCCGCCCCAACAACAACACCCGCAAGAGC
Val120-Ile201-Ile424-Ala433	(649) ACCCGCCCCAACAACAACACCCGCAAGAGC
Val120-Ile201B-Ile424-Ala433	(649) ACCCGCCCCAACAACAACACCCGCAAGAGC
Consensus	(691) ACCCGCCCCAACAACAACACCCGCAAGAGC
	721 750
Leu122-Ser199-Tryp427-Gly431	(691) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Val127-Asn195-Arg426-Gly431	(721) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Val120-Thr202-Ile424-Ala433	(679) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Leu122-Ser199-Arg426-Lys432	(691) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Leu122-Ser199-Arg426-Gly431	(691) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Lys121-Val200-Asn425-Lys432	(685) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Val120-Ile201-Ile424-Ala433	(679) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Val120-Ile201B-Ile424-Ala433	(679) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
Consensus	(721) ATCACCATCGGGCCCCGGCGCGCCTTCTAC
	751 780
Leu122-Ser199-Tryp427-Gly431	(721) GCCACCGGGGACATCATCGGCGACATCCGC
Val127-Asn195-Arg426-Gly431	(751) GCCACCGGGGACATCATCGGCGACATCCGC
Val120-Thr202-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCGACATCCGC
Leu122-Ser199-Arg426-Lys432	(721) GCCACCGGGGACATCATCGGCGACATCCGC
Leu122-Ser199-Arg426-Gly431	(721) GCCACCGGGGACATCATCGGCGACATCCGC
Lys121-Val200-Asn425-Lys432	(715) GCCACCGGGGACATCATCGGCGACATCCGC
Val120-Ile201-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCGACATCCGC
Val120-Ile201B-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCGACATCCGC
Consensus	(751) GCCACCGGGGACATCATCGGCGACATCCGC
	781 810
Leu122-Ser199-Tryp427-Gly431	(751) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Val127-Asn195-Arg426-Gly431	(781) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Val120-Thr202-Ile424-Ala433	(739) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Leu122-Ser199-Arg426-Lys432	(751) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Leu122-Ser199-Arg426-Gly431	(751) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Lys121-Val200-Asn425-Lys432	(745) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Val120-Ile201-Ile424-Ala433	(739) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Val120-Ile201B-Ile424-Ala433	(739) CAGGCCCACTGCAACATCAGCGGCGAGAAG
Consensus	(781) CAGGCCCACTGCAACATCAGCGGCGAGAAG
	811 840
Leu122-Ser199-Tryp427-Gly431	(781) TGGAAACAACCCCTGAAGCAGATCGTGACC
Val127-Asn195-Arg426-Gly431	(811) TGGAAACAACCCCTGAAGCAGATCGTGACC
Val120-Thr202-Ile424-Ala433	(769) TGGAAACAACCCCTGAAGCAGATCGTGACC
Leu122-Ser199-Arg426-Lys432	(781) TGGAAACAACCCCTGAAGCAGATCGTGACC
Leu122-Ser199-Arg426-Gly431	(781) TGGAAACAACCCCTGAAGCAGATCGTGACC
Lys121-Val200-Asn425-Lys432	(775) TGGAAACAACCCCTGAAGCAGATCGTGACC
Val120-Ile201-Ile424-Ala433	(769) TGGAAACAACCCCTGAAGCAGATCGTGACC
Val120-Ile201B-Ile424-Ala433	(769) TGGAAACAACCCCTGAAGCAGATCGTGACC
Consensus	(811) TGGAAACAACCCCTGAAGCAGATCGTGACC
	841 870
Leu122-Ser199-Tryp427-Gly431	(811) AAGCTGCAGGCCAGTTTCGGCAACAAGACC
Val127-Asn195-Arg426-Gly431	(841) AAGCTGCAGGCCAGTTTCGGCAACAAGACC
Val120-Thr202-Ile424-Ala433	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACC
Leu122-Ser199-Arg426-Lys432	(811) AAGCTGCAGGCCAGTTTCGGCAACAAGACC

FIG. 5E

Leu122-Ser199-Arg426-Gly431	(811)	AAGCTGCAGGCCAGTTCGGCAACAAGACC
Lys121-Val200-Asn425-Lys432	(805)	AAGCTGCAGGCCAGTTCGGCAACAAGACC
Val120-Ile201-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAAGACC
Val120-Ile201B-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAAGACC
Consensus	(841)	AAGCTGCAGGCCAGTTCGGCAACAAGACC
Leu122-Ser199-Tryp427-Gly431	871	900
Val127-Asn195-Arg426-Gly431	(841)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Val120-Thr202-Ile424-Ala433	(871)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Leu122-Ser199-Arg426-Lys432	(829)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Leu122-Ser199-Arg426-Gly431	(841)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Lys121-Val200-Asn425-Lys432	(841)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Val120-Ile201-Ile424-Ala433	(835)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Val120-Ile201B-Ile424-Ala433	(829)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Consensus	(829)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Leu122-Ser199-Tryp427-Gly431	(871)	ATCGTGTTCAGCAGAGCAGCGGCGGGCAG
Val127-Asn195-Arg426-Gly431	901	930
Val120-Thr202-Ile424-Ala433	(871)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Lys432	(901)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Gly431	(859)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Lys121-Val200-Asn425-Lys432	(871)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Val120-Ile201-Ile424-Ala433	(871)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Val120-Ile201B-Ile424-Ala433	(865)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Consensus	(859)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Leu122-Ser199-Tryp427-Gly431	(859)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Val127-Asn195-Arg426-Gly431	(901)	CCCGAGATCGTGTATGCACAGCTTCAACTGC
Val120-Thr202-Ile424-Ala433	931	960
Leu122-Ser199-Arg426-Lys432	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Gly431	(931)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Lys121-Val200-Asn425-Lys432	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Ile201-Ile424-Ala433	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Ile201B-Ile424-Ala433	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Consensus	(895)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Tryp427-Gly431	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val127-Asn195-Arg426-Gly431	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Thr202-Ile424-Ala433	(931)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Lys432	961	990
Leu122-Ser199-Arg426-Gly431	(931)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Lys121-Val200-Asn425-Lys432	(961)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Val120-Ile201-Ile424-Ala433	(919)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Val120-Ile201B-Ile424-Ala433	(931)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Consensus	(931)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Leu122-Ser199-Tryp427-Gly431	(931)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Val127-Asn195-Arg426-Gly431	(925)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Val120-Thr202-Ile424-Ala433	(919)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Leu122-Ser199-Arg426-Lys432	(919)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Leu122-Ser199-Arg426-Gly431	(919)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Lys121-Val200-Asn425-Lys432	(961)	CAGCTGTTCACAGCAGCCTGGAACAACACC
Val120-Ile201-Ile424-Ala433	991	1020
Val120-Ile201B-Ile424-Ala433	(961)	ATCGGCCCCAACACACCAACGGCACCATC
Consensus	(991)	ATCGGCCCCAACACACCAACGGCACCATC
Leu122-Ser199-Tryp427-Gly431	(949)	ATCGGCCCCAACACACCAACGGCACCATC
Val127-Asn195-Arg426-Gly431	(961)	ATCGGCCCCAACACACCAACGGCACCATC
Val120-Thr202-Ile424-Ala433	(961)	ATCGGCCCCAACACACCAACGGCACCATC
Leu122-Ser199-Arg426-Lys432	(961)	ATCGGCCCCAACACACCAACGGCACCATC
Leu122-Ser199-Arg426-Gly431	(955)	ATCGGCCCCAACACACCAACGGCACCATC
Lys121-Val200-Asn425-Lys432	(949)	ATCGGCCCCAACACACCAACGGCACCATC
Val120-Ile201-Ile424-Ala433	(949)	ATCGGCCCCAACACACCAACGGCACCATC
Val120-Ile201B-Ile424-Ala433	(949)	ATCGGCCCCAACACACCAACGGCACCATC
Consensus	(991)	ATCGGCCCCAACACACCAACGGCACCATC
Leu122-Ser199-Tryp427-Gly431	1021	1050
	(991)	ACCCTGCCCTGCCGCATCAAGCAGATCATC

FIG. 5F

Val127-Asn195-Arg426-Gly431	(1021)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Val120-Thr202-Ile424-Ala433	(979)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Lys432	(991)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Gly431	(991)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Lys121-Val200-Asn425-Lys432	(985)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Val120-Ile201-Ile424-Ala433	(979)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Val120-Ile201B-Ile424-Ala433	(979)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Consensus	(1021)	ACCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199 Tryp427-Gly431	(1021)	AACCGCTGGGGCGGCAAGGCCATGTACGCC
Val127-Asn195-Arg426-Gly431	(1051)	AACCGCGGGCGGGCGCAAGGCCATGTACGCC
Val120-Thr202-Ile424-Ala433	(1009)	-----GGCGGC---GCCATGTACGCC
Leu122-Ser199-Arg426-Lys432	(1021)	AACCGCGGGCGGCAACAAGGCCATGTACGCC
Leu122-Ser199-Arg426-Gly431	(1021)	AACCGCGGGCAAGGCCAAGGCCATGTACGCC
Lys121-Val200-Asn425-Lys432	(1015)	AAC-----GCCCCAAGGCCATGTACGCC
Val120-Ile201-Ile424-Ala433	(1009)	-----GGCGGC---GCCATGTACGCC
Val120-Ile201B-Ile424-Ala433	(1009)	-----GGCGGC---GCCATGTACGCC
Consensus	(1051)	AACCGC G GCGGCAAGGCCATGTACGCC
Leu122-Ser199 Tryp427-Gly431	(1051)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Val127-Asn195-Arg426-Gly431	(1081)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Val120-Thr202-Ile424-Ala433	(1027)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Lys432	(1051)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Gly431	(1051)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Lys121-Val200-Asn425-Lys432	(1039)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Val120-Ile201-Ile424-Ala433	(1027)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Val120-Ile201B-Ile424-Ala433	(1027)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Consensus	(1081)	CCCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199 Tryp427-Gly431	(1081)	AGCAACATCACC GGCTGCTGCTGACCCGC
Val127-Asn195-Arg426-Gly431	(1111)	AGCAACATCACC GGCTGCTGCTGACCCGC
Val120-Thr202-Ile424-Ala433	(1057)	AGCAACATCACC GGCTGCTGCTGACCCGC
Leu122-Ser199-Arg426-Lys432	(1081)	AGCAACATCACC GGCTGCTGCTGACCCGC
Leu122-Ser199-Arg426-Gly431	(1081)	AGCAACATCACC GGCTGCTGCTGACCCGC
Lys121-Val200-Asn425-Lys432	(1069)	AGCAACATCACC GGCTGCTGCTGACCCGC
Val120-Ile201-Ile424-Ala433	(1057)	AGCAACATCACC GGCTGCTGCTGACCCGC
Val120-Ile201B-Ile424-Ala433	(1057)	AGCAACATCACC GGCTGCTGCTGACCCGC
Consensus	(1111)	AGCAACATCACC GGCTGCTGCTGACCCGC
Leu122-Ser199 Tryp427-Gly431	(1111)	GACGGCGGCAAGGAGATCAGCAACACCACC
Val127-Asn195-Arg426-Gly431	(1141)	GACGGCGGCAAGGAGATCAGCAACACCACC
Val120-Thr202-Ile424-Ala433	(1087)	GACGGCGGCAAGGAGATCAGCAACACCACC
Leu122-Ser199-Arg426-Lys432	(1111)	GACGGCGGCAAGGAGATCAGCAACACCACC
Leu122-Ser199-Arg426-Gly431	(1111)	GACGGCGGCAAGGAGATCAGCAACACCACC
Lys121-Val200-Asn425-Lys432	(1099)	GACGGCGGCAAGGAGATCAGCAACACCACC
Val120-Ile201-Ile424-Ala433	(1087)	GACGGCGGCAAGGAGATCAGCAACACCACC
Val120-Ile201B-Ile424-Ala433	(1087)	GACGGCGGCAAGGAGATCAGCAACACCACC
Consensus	(1141)	GACGGCGGCAAGGAGATCAGCAACACCACC
Leu122-Ser199 Tryp427-Gly431	(1141)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Val127-Asn195-Arg426-Gly431	(1171)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Val120-Thr202-Ile424-Ala433	(1117)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Leu122-Ser199-Arg426-Lys432	(1141)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Leu122-Ser199-Arg426-Gly431	(1141)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Lys121-Val200-Asn425-Lys432	(1129)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Val120-Ile201-Ile424-Ala433	(1117)	GAGATCTTCGCCCCGGCGGGCGGCGACATG
Val120-Ile201B-Ile424-Ala433	(1117)	GAGATCTTCGCCCCGGCGGGCGGCGACATG

FIG. 5G

Consensus	(1171)	GAGATCTTCCGCCCGCGCGCGGCGACATG	1201	1230
Leu122-Ser199 Tryp427-Gly431	(1171)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Val127-Asn195-Arg426-Gly431	(1201)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Val120-Thr202-Ile424-Ala433	(1147)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Leu122-Ser199-Arg426-Lys432	(1171)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Leu122-Ser199-Arg426-Gly431	(1171)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Lys121-Val200-Asn425-Lys432	(1159)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Val120-Ile201-Ile424-Ala433	(1147)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Val120-Ile201B-Ile424-Ala433	(1147)	CGCGACAACCTGCGCGAGCGAGCTGTACAA		
Consensus	(1201)	CGCGACAACCTGCGCGAGCGAGCTGTACAA	1231	1260
Leu122-Ser199 Tryp427-Gly431	(1201)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Val127-Asn195-Arg426-Gly431	(1231)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Val120-Thr202-Ile424-Ala433	(1177)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Leu122-Ser199-Arg426-Lys432	(1201)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Leu122-Ser199-Arg426-Gly431	(1201)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Lys121-Val200-Asn425-Lys432	(1189)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Val120-Ile201-Ile424-Ala433	(1177)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Val120-Ile201B-Ile424-Ala433	(1177)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG		
Consensus	(1231)	TACAAGGTGGTGAAGATCGAGCCCCCTGGGG	1261	1290
Leu122-Ser199 Tryp427-Gly431	(1231)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Val127-Asn195-Arg426-Gly431	(1261)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Val120-Thr202-Ile424-Ala433	(1207)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Leu122-Ser199-Arg426-Lys432	(1231)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Leu122-Ser199-Arg426-Gly431	(1231)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Lys121-Val200-Asn425-Lys432	(1219)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Val120-Ile201-Ile424-Ala433	(1207)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Val120-Ile201B-Ile424-Ala433	(1207)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG		
Consensus	(1261)	GTGGCCCCCAACCAAGGCCAAGCGCCGCGTG	1291	1320
Leu122-Ser199 Tryp427-Gly431	(1261)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Val127-Asn195-Arg426-Gly431	(1291)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Val120-Thr202-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Leu122-Ser199-Arg426-Lys432	(1261)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Leu122-Ser199-Arg426-Gly431	(1261)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Lys121-Val200-Asn425-Lys432	(1249)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Val120-Ile201-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Val120-Ile201B-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG		
Consensus	(1291)	GTGCAGCGCGAGAGCGCGCGCGTGACCCCTG	1321	1350
Leu122-Ser199 Tryp427-Gly431	(1291)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Val127-Asn195-Arg426-Gly431	(1321)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Val120-Thr202-Ile424-Ala433	(1267)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Leu122-Ser199-Arg426-Lys432	(1291)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Leu122-Ser199-Arg426-Gly431	(1291)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Lys121-Val200-Asn425-Lys432	(1279)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Val120-Ile201-Ile424-Ala433	(1267)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Val120-Ile201B-Ile424-Ala433	(1267)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC		
Consensus	(1321)	GGCGCCATGTTCTCTGGGCTTCCTGGGCGCC	1351	1380
Leu122-Ser199 Tryp427-Gly431	(1321)	GCCGGGAGGACCATGGGGGGGCGGAGGCTG		
Val127-Asn195-Arg426-Gly431	(1351)	GCCGGGAGGACCATGGGGGGGCGGAGGCTG		
Val120-Thr202-Ile424-Ala433	(1297)	GCCGGGAGGACCATGGGGGGGCGGAGGCTG		
Leu122-Ser199-Arg426-Lys432	(1321)	GCCGGGAGGACCATGGGGGGGCGGAGGCTG		
Leu122-Ser199-Arg426-Gly431	(1321)	GCCGGGAGGACCATGGGGGGGCGGAGGCTG		

FIG. 5H

Lys121-Val200-Asn425-Lys432	(1309)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201-Ile424-Ala433	(1297)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Val120-Ile201B-Ile424-Ala433	(1297)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Consensus	(1351)	GCCGGCAGCACCATGGGCGCCCGCAGCCTG
Leu122-Ser199 Tryp427-Gly431	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Val127-Asn195-Arg426-Gly431	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Val120-Thr202-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(1351)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Lys121-Val200-Asn425-Lys432	(1339)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Val120-Ile201-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(1327)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Consensus	(1381)	ACCCTGACCGTGCAGGCCCGCCAGCTGCTG
Leu122-Ser199 Tryp427-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Val127-Asn195-Arg426-Gly431	(1411)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Val120-Thr202-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Leu122-Ser199-Arg426-Lys432	(1381)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Leu122-Ser199-Arg426-Gly431	(1381)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Lys121-Val200-Asn425-Lys432	(1369)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Val120-Ile201-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Val120-Ile201B-Ile424-Ala433	(1357)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Consensus	(1411)	AGCGGCATCGTGCAGCAGCAGAACAACCTG
Leu122-Ser199 Tryp427-Gly431	(1411)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Val127-Asn195-Arg426-Gly431	(1441)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Val120-Thr202-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Leu122-Ser199-Arg426-Lys432	(1411)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Leu122-Ser199-Arg426-Gly431	(1411)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Lys121-Val200-Asn425-Lys432	(1399)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Val120-Ile201-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Val120-Ile201B-Ile424-Ala433	(1387)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Consensus	(1441)	CTGCGCGCCATCGAGGCCCGAGCAGCACCTG
Leu122-Ser199 Tryp427-Gly431	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Val127-Asn195-Arg426-Gly431	(1471)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Val120-Thr202-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Lys432	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Leu122-Ser199-Arg426-Gly431	(1441)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Lys121-Val200-Asn425-Lys432	(1429)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Val120-Ile201-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Val120-Ile201B-Ile424-Ala433	(1417)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Consensus	(1471)	CTGCAGCTGACCGTGTGGGGCATCAAGCAG
Leu122-Ser199 Tryp427-Gly431	(1471)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Val127-Asn195-Arg426-Gly431	(1501)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Val120-Thr202-Ile424-Ala433	(1447)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Leu122-Ser199-Arg426-Lys432	(1471)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Leu122-Ser199-Arg426-Gly431	(1471)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Lys121-Val200-Asn425-Lys432	(1459)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Val120-Ile201-Ile424-Ala433	(1447)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Val120-Ile201B-Ile424-Ala433	(1447)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Consensus	(1501)	CTGCAGGCGCGCGTGTGGCGGTGGAGCGC
Leu122-Ser199 Tryp427-Gly431	(1501)	TACCTGAAGGACCAGCAGCTGCTGGGCATC
Val127-Asn195-Arg426-Gly431	(1531)	TACCTGAAGGACCAGCAGCTGCTGGGCATC

FIG. 5L

Val120-Thr202-Ile424-Ala433	(1477)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Leu122-Ser199-Arg426-Lys432	(1501)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Leu122-Ser199-Arg426-Gly431	(1501)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Lys121-Val200-Asn425-Lys432	(1489)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Val120-Ile201-Ile424-Ala433	(1477)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Val120-Ile201B-Ile424-Ala433	(1477)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Consensus	(1531)	TACCTGAAGGACCAAGCAGCTGCTGGGCATC
Leu122-Ser199 Tryp427-Gly431	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Val127-Asn195-Arg426-Gly431	(1561)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Val120-Thr202-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leu122-Ser199-Arg426-Lys432	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leu122-Ser199-Arg426-Gly431	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Lys121-Val200-Asn425-Lys432	(1519)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Val120-Ile201-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Val120-Ile201B-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Consensus	(1561)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leu122-Ser199 Tryp427-Gly431	(1561)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Val127-Asn195-Arg426-Gly431	(1591)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Val120-Thr202-Ile424-Ala433	(1537)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Leu122-Ser199-Arg426-Lys432	(1561)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Leu122-Ser199-Arg426-Gly431	(1561)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Lys121-Val200-Asn425-Lys432	(1549)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Val120-Ile201-Ile424-Ala433	(1537)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Val120-Ile201B-Ile424-Ala433	(1537)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Consensus	(1591)	ACCGCGGTGCCCTTGAACGCCAGCTGGAGC
Leu122-Ser199 Tryp427-Gly431	(1591)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Val127-Asn195-Arg426-Gly431	(1621)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Val120-Thr202-Ile424-Ala433	(1567)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Leu122-Ser199-Arg426-Lys432	(1591)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Leu122-Ser199-Arg426-Gly431	(1591)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Lys121-Val200-Asn425-Lys432	(1579)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Val120-Ile201-Ile424-Ala433	(1567)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Val120-Ile201B-Ile424-Ala433	(1567)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Consensus	(1621)	AACAAGAGCCTTGAACGAGATCTGGAACAAC
Leu122-Ser199 Tryp427-Gly431	(1621)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Val127-Asn195-Arg426-Gly431	(1651)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Val120-Thr202-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Leu122-Ser199-Arg426-Lys432	(1621)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Leu122-Ser199-Arg426-Gly431	(1621)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Lys121-Val200-Asn425-Lys432	(1609)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Val120-Ile201-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Val120-Ile201B-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Consensus	(1651)	ATGACCTGGATGGAGTGGGAGCGCGAGATC
Leu122-Ser199 Tryp427-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCCTG
Val127-Asn195-Arg426-Gly431	(1681)	GACAACTACACCAACCTGATCTACACCCTG
Val120-Thr202-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCCTG
Leu122-Ser199-Arg426-Lys432	(1651)	GACAACTACACCAACCTGATCTACACCCTG
Leu122-Ser199-Arg426-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCCTG
Lys121-Val200-Asn425-Lys432	(1639)	GACAACTACACCAACCTGATCTACACCCTG
Val120-Ile201-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCCTG
Val120-Ile201B-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCCTG
Consensus	(1681)	GACAACTACACCAACCTGATCTACACCCTG

FIG. 5J

		1711	1740
Leu122-Ser199 Tryp427-Gly431	(1681)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Val127-Asn195-Arg426-Gly431	(1711)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Val120-Thr202-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Leu122-Ser199-Arg426-Lys432	(1681)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Leu122-Ser199-Arg426-Gly431	(1681)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Lys121-Val200-Asn425-Lys432	(1669)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Val120-Ile201-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Val120-Ile201B-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
Consensus	(1711)	ATCGAGGAGAGCCAGAACCAGCAGGAGAAG	
		1741	1770
Leu122-Ser199 Tryp427-Gly431	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val127-Asn195-Arg426-Gly431	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Thr202-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Lys432	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Gly431	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Lys121-Val200-Asn425-Lys432	(1699)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Ile201-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val120-Ile201B-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Consensus	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
		1771	1800
Leu122-Ser199 Tryp427-Gly431	(1741)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Val127-Asn195-Arg426-Gly431	(1771)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Val120-Thr202-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Leu122-Ser199-Arg426-Lys432	(1741)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Leu122-Ser199-Arg426-Gly431	(1741)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Lys121-Val200-Asn425-Lys432	(1729)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Val120-Ile201-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Val120-Ile201B-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
Consensus	(1771)	TGGGCCAGCCTGTGGAACCTGGTTCCGACATC	
		1801	1830
Leu122-Ser199 Tryp427-Gly431	(1771)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val127-Asn195-Arg426-Gly431	(1801)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val120-Thr202-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Lys432	(1771)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Gly431	(1771)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Lys121-Val200-Asn425-Lys432	(1759)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val120-Ile201-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val120-Ile201B-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Consensus	(1801)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
		1831	1860
Leu122-Ser199 Tryp427-Gly431	(1801)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Val127-Asn195-Arg426-Gly431	(1831)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Val120-Thr202-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Leu122-Ser199-Arg426-Lys432	(1801)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Leu122-Ser199-Arg426-Gly431	(1801)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Lys121-Val200-Asn425-Lys432	(1789)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Val120-Ile201-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Val120-Ile201B-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
Consensus	(1831)	ATCATGATCGTGGGCGGCCTGGTGGGCCTG	
		1861	1890
Leu122-Ser199 Tryp427-Gly431	(1831)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	
Val127-Asn195-Arg426-Gly431	(1861)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	
Val120-Thr202-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Lys432	(1831)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Gly431	(1831)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	
Lys121-Val200-Asn425-Lys432	(1819)	CGCATCGTGTTCACCGTGCTGAGCATCGTG	

FIG. 5K

Vall120-Ile201-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Vall120-Ile201B-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Consensus	(1861)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Leu122-Ser199 Tryp427-Gly431	(1861)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall127-Asn195-Arg426-Gly431	(1891)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall120-Thr202-Ile424-Ala433	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Lys432	(1861)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Gly431	(1861)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Lys121-Val200-Asn425-Lys432	(1849)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall120-Ile201-Ile424-Ala433	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall120-Ile201B-Ile424-Ala433	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Consensus	(1891)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199 Tryp427-Gly431	(1891)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Vall127-Asn195-Arg426-Gly431	(1921)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Vall120-Thr202-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Leu122-Ser199-Arg426-Lys432	(1891)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Leu122-Ser199-Arg426-Gly431	(1891)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Lys121-Val200-Asn425-Lys432	(1879)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Vall120-Ile201-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Vall120-Ile201B-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Consensus	(1921)	AGCTTCCAGACCCGCTTCCCGCCCCCGC
Leu122-Ser199 Tryp427-Gly431	(1921)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Vall127-Asn195-Arg426-Gly431	(1951)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Vall120-Thr202-Ile424-Ala433	(1897)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Leu122-Ser199-Arg426-Lys432	(1921)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Leu122-Ser199-Arg426-Gly431	(1921)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Lys121-Val200-Asn425-Lys432	(1909)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Vall120-Ile201-Ile424-Ala433	(1897)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Vall120-Ile201B-Ile424-Ala433	(1897)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Consensus	(1951)	GGCCCCGACCGCCCGAGGCGATCGAGGAG
Leu122-Ser199 Tryp427-Gly431	(1951)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Vall127-Asn195-Arg426-Gly431	(1981)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Vall120-Thr202-Ile424-Ala433	(1927)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Leu122-Ser199-Arg426-Lys432	(1951)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Leu122-Ser199-Arg426-Gly431	(1951)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Lys121-Val200-Asn425-Lys432	(1939)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Vall120-Ile201-Ile424-Ala433	(1927)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Vall120-Ile201B-Ile424-Ala433	(1927)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Consensus	(1981)	GAGGGCGGGCGAGCGCGACCGCGACCGCAGC
Leu122-Ser199 Tryp427-Gly431	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Vall127-Asn195-Arg426-Gly431	(2011)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Vall120-Thr202-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Lys432	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199-Arg426-Gly431	(1981)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Lys121-Val200-Asn425-Lys432	(1969)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Vall120-Ile201-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Vall120-Ile201B-Ile424-Ala433	(1957)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Consensus	(2011)	AGCCCCCTGGTGCACGGCCTGCTGGCCCTG
Leu122-Ser199 Tryp427-Gly431	(2011)	ATCTGGGAGGAGCGCGGAGCGGAGCGGAGC
Vall127-Asn195-Arg426-Gly431	(2041)	ATCTGGGAGGAGCGCGGAGCGGAGCGGAGC
Vall120-Thr202-Ile424-Ala433	(1987)	ATCTGGGAGGAGCGCGGAGCGGAGCGGAGC

FIG. 5L

Leu122-Ser199-Arg426-Lys432	(2011)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Leu122-Ser199-Arg426-Gly431	(2011)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Lys121-Val200-Asn425-Lys432	(1999)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Val120-Ile201-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Val120-Ile201B-Ile424-Ala433	(1987)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Consensus	(2041)	ATCTGGGACGACCTGCGCAGCCTGTGCCTG
Leu122-Ser199 Tryp427-Gly431	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val127-Asn195-Arg426-Gly431	(2071)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Thr202-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Leu122-Ser199-Arg426-Lys432	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Leu122-Ser199-Arg426-Gly431	(2041)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Lys121-Val200-Asn425-Lys432	(2029)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Ile201-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Val120-Ile201B-Ile424-Ala433	(2017)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Consensus	(2071)	TTCAGCTACCACCGCCTGCGCGACCTGATC
Leu122-Ser199 Tryp427-Gly431	(2071)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Val127-Asn195-Arg426-Gly431	(2101)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Val120-Thr202-Ile424-Ala433	(2047)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(2071)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(2071)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Lys121-Val200-Asn425-Lys432	(2059)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Val120-Ile201-Ile424-Ala433	(2047)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(2047)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Consensus	(2101)	CTGATCGCGCGCGCGCATCGTGGAGCTGCTG
Leu122-Ser199 Tryp427-Gly431	(2101)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val127-Asn195-Arg426-Gly431	(2131)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Thr202-Ile424-Ala433	(2077)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Lys432	(2101)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Gly431	(2101)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Lys121-Val200-Asn425-Lys432	(2089)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Ile201-Ile424-Ala433	(2077)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Ile201B-Ile424-Ala433	(2077)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Consensus	(2131)	GGCGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199 Tryp427-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val127-Asn195-Arg426-Gly431	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Thr202-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Lys432	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Leu122-Ser199-Arg426-Gly431	(2131)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Lys121-Val200-Asn425-Lys432	(2119)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Ile201-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Val120-Ile201B-Ile424-Ala433	(2107)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Consensus	(2161)	TGGGGCAACCTGCTGCAGTACTGGATCCAG
Leu122-Ser199 Tryp427-Gly431	(2161)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Val127-Asn195-Arg426-Gly431	(2191)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Val120-Thr202-Ile424-Ala433	(2137)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Leu122-Ser199-Arg426-Lys432	(2161)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Leu122-Ser199-Arg426-Gly431	(2161)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Lys121-Val200-Asn425-Lys432	(2149)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Val120-Ile201-Ile424-Ala433	(2137)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Val120-Ile201B-Ile424-Ala433	(2137)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
Consensus	(2191)	GAGCTGAAGAACAGCGCCCTGAGCCTGTTT
	2221	2250

FIG. 5M

Leu122-Ser199 Tryp427-Gly431	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val127-Asn195-Arg426-Gly431	(2221)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val120-Thr202-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Lys432	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Gly431	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Lys121-Val200-Asn425-Lys432	(2179)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val120-Ile201-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val120-Ile201B-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Consensus	(2221)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199 Tryp427-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Val127-Asn195-Arg426-Gly431	(2251)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Val120-Thr202-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Leu122-Ser199-Arg426-Lys432	(2221)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Leu122-Ser199-Arg426-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Lys121-Val200-Asn425-Lys432	(2209)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Val120-Ile201-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Val120-Ile201B-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Consensus	(2251)	ACCGACCGCATCATCGAGGTGGCCAGGCGC
Leu122-Ser199 Tryp427-Gly431	(2251)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Val127-Asn195-Arg426-Gly431	(2281)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Val120-Thr202-Ile424-Ala433	(2227)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Leu122-Ser199-Arg426-Lys432	(2251)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Leu122-Ser199-Arg426-Gly431	(2251)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Lys121-Val200-Asn425-Lys432	(2239)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Val120-Ile201-Ile424-Ala433	(2227)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Val120-Ile201B-Ile424-Ala433	(2227)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Consensus	(2281)	ATCGGCGCGCGCTTCCTGCACATCCCCCGC
Leu122-Ser199 Tryp427-Gly431	(2281)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Val127-Asn195-Arg426-Gly431	(2311)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Val120-Thr202-Ile424-Ala433	(2257)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199-Arg426-Lys432	(2281)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199-Arg426-Gly431	(2281)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Lys121-Val200-Asn425-Lys432	(2269)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Val120-Ile201-Ile424-Ala433	(2257)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Val120-Ile201B-Ile424-Ala433	(2257)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Consensus	(2311)	CGCATCCGCCAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199 Tryp427-Gly431	(2311)	CTGTAAGTCGAG
Val127-Asn195-Arg426-Gly431	(2341)	CTGTAAGTCGAG
Val120-Thr202-Ile424-Ala433	(2287)	CTGTAAGTCGAG
Leu122-Ser199-Arg426-Lys432	(2311)	CTGTAAGTCGAG
Leu122-Ser199-Arg426-Gly431	(2311)	CTGTAAGTCGAG
Lys121-Val200-Asn425-Lys432	(2299)	CTGTAAGTCGAG
Val120-Ile201-Ile424-Ala433	(2287)	CTGTAAGTCGAG
Val120-Ile201B-Ile424-Ala433	(2287)	CTGTAAGTCGAG
Consensus	(2341)	CTGTAAGTCGAG

FIG. 5N

SEQ ID NO:3 VAL120-ALA204

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGCCGGCGCCTGCCCCAA
GGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTG
CAACGACAAGAAGTTCAACGGCAGCGGGCCCTGCACCAACGTGAGCACCGTGCAAGTGCACCC
ACGGCATCCGCCCCGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGC
GTGGTGATCCGCAGCGAGAACTTACCCGACAACGCCAAGACCATCATCGTGCAAGTGAAGGA
GAGCGTGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCC
CCGGCCGCGCCTTCTACGCCACCGGGCAGCATCATCGGCGACATCCGCCAGGCCCACTGCAACA
TCAGCGGCGAGAAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTC
GGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAG
CTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAA
CAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGA
TCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATC
CGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAA
CACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGT
ACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGCCCCCAACCAAGGCCAAGCGCCGC
GTGGTGACGCGGAGAAGCGCGCCGTGACCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCC
GCCGGCAGCACCATGGGCGCCCCGAGCCTGACCTGACCGTGACGGCCCGCCAGCTGCTGAG
CGGCATCGTGACGAGCAGAAACACCTGCTGCGCGCCATCGAGGCCCGAGCAGCACCTGCTGC
AGCTGACCGTGTTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTG
AAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAACCGCCGT
GCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGA
TGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCTACACCTGATCGAGGAGAGC
CAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGT
GGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCG
GCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCT
ACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCA
TCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTG
GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTTACGCTACCACCGCCTGCGCGACCTG
ATCCTGATCGCCGCCCCGATCGTGGAGCTGCTGGGCCGCGCGGCTGGGAGGCCCTGAAGTAC
TGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCA
CGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCG
GCCGCGCCTTCCTGCACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAAC
TCGAG

FIG. 6

SEQ ID NO:4 VAL120-ILE201

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCATCACCCAGGCCTG
CCCCAAGGTGAGCTTCGAGGCCATCCCCATCCACTACTGCGCCCCCGCGGGCTTCGCCATCCT
GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGT
GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAG
GAGGGCGTGGTGATCCGCAGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGCAGCT
GAAGGAGAGCGTGAGATCAACTGCACCCGCCCCAACAAACAACACCCGCAAGAGCATCACCA
TCGGCCCCGGCGCGCCTTCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCCACT
GCAACATCAGCGCGGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC
CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGCGCGGACCCCGAGATCGTGAT
GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC
CTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCA
AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCATCCGCGGC
CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGAT
CAGCAACACCACCGAGATCTTCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCG
AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAG
CGCCGCGTGGTGACGCGGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTG
GGCGCCCGCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCCAGCT
GCTGAGCGGCATCGTGACGAGCAGACAACCTGCTGCGCGCCATCGAGGCCCGAGCAGCACC
TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGACGGCCCGCGTGTGGCCGTGGAGCGC
TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC
CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA
CCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCTACACCCTGATCGAG
GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA
GCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCC
AGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCG
AGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGG
CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTACAGTACCACCGCCTGOG
CGACCTGATCCTGATCGCCGCCCCGATCGTGAGCTGCTGGGCCGCGCGGCTGGGAGGCCCT
GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC
TGTTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGC
GCATCGGCCGCGCCTTCTGACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC
TGTAACCTCGAG

FIG. 7

SEQ ID NO:5 VAL120-ILE201B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCG
TTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCA
CCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCC
ACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACCTTCAACA
TGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGC
CCTGCGTGCCCGGCATCACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGC
CCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGT
GAGCACCGTGCAGTGCAACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCT
GGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGACGCT
GAAGGAGAGCGTGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCC
CGGCCGCGCCTTCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCAGCGGC
GAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGACCATC
CTGTTCAAGCAGAGCAGCGCGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTC
TTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCCCAACAACACCAAC
GGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATG
TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGAGC
GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGC
GCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCCTGGGCGTGGCCCCCAACAAGGCCAAGC
GCCGCGTGGTGACGCGGAGAAAGCGCGCGTGACCCTGGGCGCCATGTTCTTGGGCTTCTTGGGCGCCGC
CGGCAGCACCATGGGCGCCCCGACGCTGACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGT
GCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG
CATCAAGCAGCTGCAGGCCCCGCTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCAT
CTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG
CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCT
GATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGG
ACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT
GATCGTGGGCGGCTGGTGGGCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGCGCCAG
GGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGGACCGCCCCGAGGGCATCG
AGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCT
GGGACGACCTGCGCAGCCTGTGCCTGTTACGCTACCAACCGCCTGCOCGACCTGATCCTGATCGCCGCCC
CATCGTGGAGCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTG
GATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATCGCCATCGCCGTGGCCGAGGGCAC
CGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCTGACATCCCCCGCCGCATCCGCCAG
GGCTTCGAGCGCGCCCTGCTGTAACCTCGAGCGTGCT

FIG. 8

SEQ ID NO:6 LYS121-VAL200

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCACGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCGCGCTGATCAGCCA
GGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGC
CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCG
TGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCAGCCAGCTGCTGCTGAACGGCAGCCTGG
CCGAGGAGGGCGTGGTGATCCGCGAGCGAGAACTTACCGGACAACGCCAAGACCATCATCGTG
CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT
CACCATCGGCCCGCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC
CCACTGCAACATCAGCGGCGAGAAGTGGAAACAACACCTGAAGCAGATCGTGACCAAGCTGC
AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATC
GTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC
AGCACCTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGCCCTGCCG
CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCATCC
GCGGCCAGATCCGCTGCAGCAGCAACATCACCGCCTGCTGCTGACCCGCGACGGCGGCAAG
GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCG
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGCCCCCAACAAG
CCAAGCGCCGCGTGGTGACGCGGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTTGGGC
TTCTTGGGCGCCCGCGGCGAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCG
CAGCTGCTGAGCGGCATCGTGACGAGCAGAAACAACCTGCTGCGCGCCATCGAGGCCAGCA
GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGG
AGCGCTACCTGAAGGACCAGCAGCTGTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA
CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACACCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAAGCAGCAGGAGCTGCTGGAGCTGGACAAGTG
GGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT
GATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGT
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCG
CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGC
ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTTACGCTACCAACGOC
TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGAGCTGCTGGGCCCGCGCGGCTGGGAGG
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG
AGCCTGTTTCAGCGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
CAGCGCATCGGCCGCGCCTTCTGCACATCCCCCGCCCATCCGCCAGGGCTTCGAGCGCGCC
CTGCTGTAACCTCGAGCGTGCT

FIG. 9

SEQ ID NO:7: LEU122-SER199

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGTGCCCCGTG
TGGAAGGAGGCCACCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCAACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT
CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCCATCCCCATCCACTACTGCGCCCCCGCCGG
CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGA
GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCAGCTGCTGCTGAACGGC
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT
CATCGTGCACTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCA
AGAGCATCACCATCGGCCCGCGCCGCTTCTACGCCACCGGCGACATCATCGGCGACATCC
GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACC
AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC
CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT
GTTCAACAGCACCTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGC
CCTGCCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCC
CCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGC
GGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAA
CTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAGATCGAGCCCTGGGCGTGGCCCCCA
CCAAGGCCAAGCGCCGCGTGGTGAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTT
CTGGGCTTCCTGGGCGCCGCCGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGCA
GCCCGCCAGCTGCTGAGCGGCATCGTGAGCAGCAGACAACCTGCTGCGCGCCATCGAGGC
CCAGCAGCACTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGG
CCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTG
ATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTG
GAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACA
CCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGA
CAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTT
CATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAA
CCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCC
CGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCC
CTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTAGCTAC
CACCGCCTGCGCGACCTGATCCTGATCGCCGCCCCGCATCGTGAGGCTGCTGGGCGCGCGGC
TGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAG
CGCCGTGAGCCTGTTGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGA
GGTGGCCCAGCGCATCGGCCGCGCCTTCTGCACATCCCCCGCCGATCCGCCAGGGCTTCGA
GCGCGCCCTGCTGTAACTCGAGCGTGCT

FIG. 10

SEQ ID NO:8 VAL120-THR202

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG
TGGAAGGAGGCCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG
CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGQCATCCT
GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGACGT
GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAG
GAGGGCGTGGTGATCCGCGAGCGAGAATTCAACGACAACGCCAAGACCATCATCGTGACGCT
GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA
TCGGCCCCGGCGCGCCTTCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCACT
GCAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCC
CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT
GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC
CTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCA
AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCATCCGCGGC
CAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGAT
CAGCAACACCAACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCG
AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAG
CGCCGCGTGGTGACGCGGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTTGGGCTTCCTG
GGCGCCGCCGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCCAGCT
GCTGAGCGGCATCGTGACGAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCGAGCAGCACC
TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGC
TACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAC
CGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGA
CCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCTACACCCTGATCGAG
GAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA
GCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCC
AGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCG
AGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGG
CCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTACAGCTACCACCGCCTGCG
CGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCGCCGCGGGCTGGGAGGCCCT
GAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCC
TGTTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGC
GCATCGGCCGCGCCTTCTGACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGC
TGTAACCTCGAG

FIG. 11

SEQ ID NO:9 TRP427-GLY431

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAAGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATACCCAGGCCTGCCCAAGGTGAGCTTCGA
GCCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGCAAGTGCACCCACGGCATCCGCC
CCGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGCTGATCCGC
AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGACGTGAAGGAGAGCGTGAGAT
CAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCGCGCCGCGCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACCAACGGCACCATCACCTGCCCTGCCGATCAAGCAGATCATCAACCGCT
GGGGCGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATC
ACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCG
CCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGA
AGATCGAGCCCCTGGGCGTGCCCCCACCAGGCCAAGCGCCGCGTGCTGAGCGCGGAGAA
CGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCTGGGCGCCGCGGCGAGCACCATGGGC
GCCCCGAGCCTGACCCTGACCGTGAGGCCCCGCCAGCTGCTGAGCGGCATCGTGAGCAGCA
GAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGTGTTGGGGCA
TCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG
GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTG
GAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAG
ATCGACAACCTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAA
GAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCA
GCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCA
TCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC
AGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGC
GAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGGACGA
CCTGCGCAGCCTGTGCCTGTTAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCG
CATCGTGGAGCTGCTGGGCGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGC
AGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGCC
GTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCTGCA
CATCCCCGCGCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 12

SEQ ID NO:10 ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCAGGCCTGCCCCAAGGTGAGCTTCGA
GCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGGCCCTGCACCAACGTGAGCACCGTGCAAGTGACCCACGGCATCCGCC
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACCTTACCGACAACGCCAAGACCATCATCGTGCAAGTGAAGGAGAGCGTGAGAGAT
CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCGGGCCGCGCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCAACCGC
GGCGGCGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT
CACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCC
GCCCCGCGGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG
AAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGACGCGCGAGAA
GCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCCGCCGCGCAGCACCATGGG
CGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGACGAGC
AGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC
ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT
GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT
GGAGCAACAAGAGCCTGGACCATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGA
GATCGACAACCTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCCAGCAGGAGA
AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATC
AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC
ATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC
CAGACCCGCTTCCCCGCCCCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGG
CGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACG
ACCTGCGCAGCCTGTGCCTGTTTACGCTACCAACCGCCTGCGCGACCTGATCCTGATCGCCGCC
GCATCGTGAGCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGC
CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCCTGC
ACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

FIG. 13

SEQ ID NO:11 ARG426-GLY431B

GAATTTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCAAGGTGAGCTTCGA
GCCCCATCCCCATCCACTACTGCGCCCCCGCGGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCACTGCACCCACGGCATCCGCC
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGACGTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCCAACAAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCGCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCCGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAAACCAACGGCACCATCACCTGCCCTGCCGATCAAGCAGATCATCAACCGC
GGCAGCGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT
CACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCC
GCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG
AAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGACGCGGAGAA
GCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCCGCGGCGAGCACCATGGG
CGCCCGCAGCCTGACCCTGACCGTGACGGCCCCGCCAGCTGCTGAGCGGCATCGTGACGAGC
AGAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC
ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT
GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT
GGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGA
GATCGACAATAACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA
AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATC
AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC
ATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC
CAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGG
CGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACG
ACCTGCGCAGCCTGTGCCTGTTCACTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCC
GCATCGTGAGCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGC
CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGC
ACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 14

SEQ ID NO:12 ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCAACGCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCAGGCCTGCCCAAGGTGAGCTTCGA
GCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGCACTGCACTGCACTGCACTGCACT
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGACGTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCCGCGCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCAACCGC
GGCGGCAACAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT
CACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCC
GCCCCGCGGCGGCGGACATGCGCGACAAGTGCGCGAGCGAGCTGTACAAGTACAAGGTGGTG
AAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGACGCGCGAGAA
GCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCCGCGCAGCACCATGGG
CGCCCGCAGCCTGACCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGACGAGC
AGAACAACTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGC
ATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCT
GGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCT
GGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGA
GATCGACAACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGA
AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAAGTGGTTCGACATC
AGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGC
ATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTC
CAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGG
CGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACG
ACCTGCGCAGCCTGTGCCTGTTCAAGTACCAACCGCCTGCGCGACCTGATCCTGATCGCCGCC
GCATCGTGAGCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGC
CGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCCTGC
ACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 15

SEQ ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAGGAGGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCAGGCCTGCCCAAGGTGAGCTTCGA
GCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCG
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAATTACCGACAACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTCCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCAACGCC
CCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCC
TGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGC
GGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGA
GCCCCTGGGCGTGCGCCCCACCAAGGCCAAGCGCCGCGTGGTGACGCGGAGAAGCGCGCCG
TGACCCTGGGCGCCATGTTCTGCGCTTCTGCGCGCCGCGCCGCGCAGCACCATGGGCGCCCGCA
GCCTGACCCTGACCGTGACGGCCCCGACGCTGCTGAGCGGCATCGTGACGAGCAGACAACAAC
CTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCA
GCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCT
GGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAAC
AAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAA
CTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGC
AGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTTCGACATCAGCAAGTGG
CTGTGGTACATCAAGATCTTCATCATGATCGTGCGCGGCCCTGGTGGGCCTGCGCATCGTGTT
ACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGC
TTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGA
CCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAG
CCTGTGCCTGTTCACTACCAACCGCTGCGCGACCTGATCCTGATCGCCGCCCCGCATCGTGGA
GCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGA
TCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGCCGTGGCCGAG
GGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCTGCACATCCCCCG
CGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

FIG. 16

SEQ ID NO:14 ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGGTGCCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA
GCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCCCTGCACCAACGTGAGCACCGTGCAAGTGCAACCCACGGCATCCGCC
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGTTGATCCGC
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGACGTGAAGGAGAGCGTGAGAT
CAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGGCGCGCCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCGGCGGC
GCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTG
CTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGG
CGACATGCGCGACAACCTGGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCC
TGGGCGTGCCCCCACCAGGCCAAGCGCCGCGTGGTGACGCGGAGAAGCGCGCCGTGACC
CTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCCGCGCAGCACCATGGGCGCCCCGACGCTG
ACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGACGAGCAGACAACACCTGCT
GCGCGCCATCGAGGCCCGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGC
AGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC
TGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAG
CCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACA
CCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGA
GCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTTCGACATCAGCAAGTGGCTGT
GGTACATCAAGATCTTCATCATGATCGTGCGCGGCCTGGTGGGCCTGCGCATCGTGTTACCG
TGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCC
CCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGC
GACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG
TGCCTGTTTCAGTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGAGCTG
CTGGGCGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCA
GGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCA
CCGACCGCATCATCGAGGTGGCCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA
TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 17

SEQ ID NO:15 ILE423-MET434

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGGTGCCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAAGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA
GCCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCACTGCACCCACGGCATCCGCC
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCAACATCGGCCCCGCGCGCCT
TCTACGCCACCGCGACATCATCGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCGGCGGCATG
TACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACC
CGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGCGGCGGCGACAT
GCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCCTGGGCG
TGGCCCCCACCAAGGCCAAGCGCCGCTGGTGCAGCGCGAGAAGCGCGCCGTGACCTGGGC
GCCATGTTCTTGGGCTTCTTGGGCGCCGCCGCGCAGCACCATGGGCGCCCCGAGCCTGACCTG
ACCGTGCAAGGCCCGCCAGCTGCTGAGCGGCATCGTGAGCAGCAGAACAACTGCTGCGCGC
CATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCC
GCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGC
GGCAAGCTGATCTGCACCACCGCCCTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGA
CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACC
TGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTG
GAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACAT
CAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAG
CATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCC
CCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGC
AGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTG
TTCAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCCGCATCGTGGAGCTGCTGGGC
CGCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCT
GAAGAACAGCGCCGTGAGCCTGTTTCAGCGCCATCGCCATCGCCGTGGCCGAGGGCACCGACC
GCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCTGACATCCCCCGCCGCATCCGCC
AGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 18

SEQ ID NO:16 GLN422-TYR435

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCACGCGCCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
ACCCTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCACCCAGGCCTGCCCCAAGGTGAGCTTCGA
GCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAAGTGACCCACGGCATCCGCC
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGCAAGTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCAACAACAACCCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCT
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG
AAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTCCGGCAACAAGAC
CATCGTGTTCAGCAGAGCAGCGGCGGACGCCAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGGGCGGGCTACGCC
CCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGAC
GGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGA
CAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCC
CCACCAAGGCCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATG
TTCCTGGGCTTCCTGGGCGCCGCGGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTG
CAGGCCCCGCGAGCTGCTGAGCGGCATCGTGACGACGAGAACAACCTGCTGCGCGCCATCGA
GGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGC
TGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAG
CTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGAT
CTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCT
ACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCT
GGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA
TCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCG
TGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCG
GCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAG
CCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
CTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCGCCG
CGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGA
ACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATC
ATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGC
TTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 19

SEQ ID NO:17 GLN422-TYR435B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCTGCGTGCCCAACGACCCCAACCCCAAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTG
ACCTTGCACTGCACCAACCTGAAGAACGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAGC
TACAAGCTGATCAACTGCAACACCAGCGTGATCAACCAGGCCTGCCCCAAGGTGAGCTTCGA
GCCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAA
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGCAAGTGACCCACGGCATCCGCC
CCGTGGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC
AGCGAGAACTTACCGACAACGCCAAGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGAT
CAACTGCACCCGCCCCAACAACAACCCCGCAAGAGCATCAACATCGGCCCGCGCGCGCT
TCTACGCCACCGCGCATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGCGGAG
AAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC
CATCGTGTTCAAGCAGAGCAGCGGCGGACCCCGAGATCGTGATGCACAGCTTCAACTGCG
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCG
GCCCCAACAACACCAACGGCACCATCAACCTGCCCTGCCGCATCAAGCAGGCCCTACGCCC
CCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACG
GCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGAC
AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCC
CACCAGGCCAAGCGCCGCTGGTGACGCGGAGAAGCGCGCCGTGACCTGGGCGCCATGT
TCCTGGGCTTCCTGGGCGCCGCCGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGC
AGGCCCCGCCAGCTGCTGAGCGGCATCGTGACGAGCAGAACAACTGCTGCGCGCCATCGAG
GCCCAGCAGCACCTGCTGACGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGT
GGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGC
TGATCTGCACCAACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATC
TGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTA
CACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTG
GACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGAT
CTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGTGAGCATCGT
GAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGG
CCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGGACCGCGACCGCAGCAGC
CCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTACG
TACCACCGCCTGCGCGACCTGATCCTGATCGCCCGCCGATCGTGAGCTGCTGGGCGCCGC
GGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAA
CAGCGCGGTGAGCCTGTTGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCAT
CGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCTGCACATCCCCGCGCATCCGCCAGGGCTT
CGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 20

SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT
CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGG
CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGA
GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCAGCTGCTGCTGAACGGC
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAAGTTCAACGACAACGCCAAGACCAT
CATCGTGACGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA
AGAGCATCACCATCGGCCCGGGCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCC
GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACC
AAGCTGCAGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCC
CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT
GTTCAACAGCACCTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGC
CCTGCCGATCAAGCAGATCATCAACCGCGGCGGCGGAAGGCCATGTACGCCCCCCCCATCC
CGGCGCAGATCCGCTGCAGCAGCAACATCACCGGCGCTGCTGCTGACCCGCGACGGCGGCAAG
GAGATCAGCAACACCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAAGTGGCG
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG
CCAAGCGCCGCGTGGTGACGCGGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTTGGGC
TTCCTGGGCGCCGCGGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCGCG
CAGCTGCTGAGCGGCATCGTGACGAGCAGACAACCTGCTGCGCGCCATCGAGGCCAGCA
GCACCTGCTGACGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG
AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC
ACCACCGCGCTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA
CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCTACACCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG
GGCCAGCCTGTGGAAGTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT
GATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGT
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCG
CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGGACCGCGACCGCAGCAGCCCCCTGGTGC
ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTACGTAACACCGCC
TGCGCGACCTGATCCTGATCGCCGCCCCGATCGTGAGCTGCTGGGCGCGCGGCTGGGAGG
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG
AGCCTGTTGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
CAGCGCATCGGCCGCGCCTTCTGACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCC
CTGCTGTAACCTCGAG

FIG. 21

SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT
CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGG
CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCCCTGCACCAACGTGA
GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGC
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAAGTTCAACGACAACGCCAAGACCAT
CATCGTGACGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA
AGAGCATCACCATCGCCCCCGCCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCC
GCCAGGCCCACTGCAACATCAGCGGCGAGAAGTGGAAACAACACCTGAAGCAGATCGTGACC
AAGCTGCAGGCCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGCGACCC
CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT
GTTCAACAGCACCTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCCCTGC
CCTGCCGCATCAAGCAGATCATCAACCGCGGCGGCAACAAGGCCATGTACGCCCCCCCCATCC
GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG
GAGATCAGCAACACCACCGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAAGTGGCG
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGG
CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTTGGGC
TTCCTGGGCGCCGCGGCGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGC
CAGCTGCTGAGCGGCATCGTGACGAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCA
GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCCGCGTGCTGGCCGTGG
AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA
CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCTACACCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG
GGCAGCCTGTGGAAGTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT
GATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGT
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGCCCCGACCG
CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGC
ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTACGCTACCACCGCC
TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGAGCTGCTGGGCCCGCCGCGCTGGGAGG
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG
AGCCTGTTGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC
CAGCGCATCGGCCGCGCCTTCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCC
CTGCTGTAACTCGAG

FIG. 22

SEQ ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCAACGCCAACCCCAAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAAGCTGGGCAACAGCGTGAT
CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGG
CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGA
GCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGAGCACCAGCTGCTGCTGAACGGC
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCAT
CATCGTGCACTGAAGGAGAGCGTGAGATCAACTGCACCCGCCCAACAACAACACCCGCA
AGAGCATCACCATCGGCCCGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCC
GCCAGGCCCACGTGCAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACC
AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCCGACCC
CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCT
GTTCAACAGCACCTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGC
CCTGCCGCATCAAGCAGATCATCAACCGCTGGGGCGGCAAGGCCATGTACGCCCCCCCCATCC
GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAG
GAGATCAGCAACACCACCGAGATCTTCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCG
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGCCCCCAACAAGG
CCAAGCGCCGCGTGGTGCAGCGCGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTGGGC
TTCCTGGGCGCCGCGGCAGCACCATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCCG
CAGCTGCTGAGCGGCATCGTGACGACGAGAACAACCTGCTGCGCGCCATCGAGGCCCAAGCA
GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGG
AGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGC
ACCAACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAA
CATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACACCCCTGA
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG
GGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT
GATCGTGGGCGGCCTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGT
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCG
CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGAGCCCCCTGGTG
ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTAGCTACCAACGCC
TGCGCGACCTGATCCTGATCGCCGCCCGCATCGTGAGCTGCTGGGCGCCCGCGGCTGGGAGG
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG
AGCCTGTTTCGACGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGC
CAGCGCATCGGCCGCGCCTTCTGACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCC
CTGCTGTAACTCGAG

FIG. 23

SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGGCCCCCGTGATCACCCA
GGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGC
CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCG
TGCAGTGACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGTGTGAACGGCAGCCTGG
CCGAGGAGGGCGTGTTGATCCGCAGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTG
CAGCTGAAGGAGAGCGTGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCAT
CACCATCGGCCCCGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGC
CCACTGCAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGC
AGGCCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGACCCCGAGATC
GTGATGCACAGCTTCAACTGCGGCGGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC
AGCACTGGAACAACACCATCGGCCCCAACAACACCAACGGCACCATCACCTGCCCTGCCG
CATCAACGAGATCATCAACGCCCCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCG
CTGCAGCAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACA
CCACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTAC
AAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGCCCCCACCAGGCCAAGCGCCGCGT
GGTGACGCGGAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCCGC
CGGCAGCACCATGGGCGCCCCGAGCCTGACCCTGACCGTGACGGCCCCGCGAGCTGCTGAGCG
GCATCGTGACGAGCAGACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAG
CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAA
GGACCAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGC
CCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATG
GAGTGGGAGCGGAGATCGACAACCTACACCAACCTGATCTACACCTGATCGAGGAGAGCCA
GAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGG
AACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC
CTGGTGGGCCTGCGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTAC
AGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCATC
GAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGC
CCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCACTACACCGCCTGCGCGACCTGAT
CCTGATCGCCGCCCGCATCGTGGAGCTGCTGGGCGCCGCGGCTGGGAGGCCCTGAAGTACTG
GGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTCCGACG
CCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGC
CGCGCCTTCCTGCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTC
GAG

FIG. 24

SEQ ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGGTGCCCCGTG
TGGAAGGAGGGCCACCACCACTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCGCGGCATCACCCAGGCCTG
CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCT
GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGGCCCTGCACCAACGTGAGCACCCTGCAGT
GCACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGTAACGGCAGCCTGGCCGAG
GAGGGCGTGATCCGCGAGCGAGAAGTTCAACCGACAACGCCAAGACCATCATCGTGCAGCT
GAAGGAGAGCGTGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA
TCGGCCCCGGCCGCGCCTTCTACGCCACCGCGCATCATCGGCGACATCCGCCAGGCCCACT
GCAACATCAGCGCGAGAAGTGGAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCC
CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGAT
GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC
CTGGAACAACACCATCGGCCCAACAACAACCAACGGCACCATCACCTGCCCTGCCGCATCA
AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC
AACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGAT
CTTCCGCCCCGGCGGCGGCGACATGCGCGACAAGTGCGCGAGCGAGCTGTACAAGTACAAGG
TGGTGAAGATCGAGCCCCCTGGGCGTGCCCCCAACAAGGCCAAGCGCCGCGTGCTGCAGCGC
GAGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCTGGGCGCCGCGGCAGCACC
ATGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGCA
GCAGCAGAACAACTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGT
GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG
CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC
AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCG
CGAGATCGACAATAACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGG
AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAAGTGGTTCGAC
ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTG
CGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC
TTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG
CGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG
ACGACCTGCGCAGCCTGTGCCTGTTAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCG
CCCGCATCGTGGAGCTGCTGGGCGCGCGCGGTGGGAGGCCCTGAAGTACTGGGGCAACCTG
CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATC
GCCGTGGCCGAGGGGACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCT
GCACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 25

SEQ ID NO:23: VAL120-ILE201B; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTTCGCCACGCGCCGTGGAGAAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGCCCGGCATCACCCAGGCCTGC
CCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTG
AAGTGCAACGACAAGAAGTTCAACGGCAGCGGGCCCTGCACCAACGTGAGCACCGTGCAAGT
CACCCACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGG
AGGGCGTGGTGATCCGCAGCGAGAACTTACCGACAACGCCAAGACCATCATCGTGCAGCTG
AAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCAT
CGGCCCCGGCCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTG
CAACATCAGCGGCGAGAAGTGGAACAACACCCTGAAGCAGATCGTGACCAAGCTGCAGGCC
AGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATG
CACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC
TGGAACAACACCATCGGCCCCAACAACAACAACGGCACCATCACCTGCCCTGCCGCATCAA
GCAGATCATCGGCGGCGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
ACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATC
TTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGT
GGTGAAGATCGAGCCCCCTGGGCGTGCCCCCAACAAGGCCAAGCGCCGCGTGGTGACGCGCG
AGAAGCGCGCCGTGACCCTGGGCGCCATGTTCTCTGGGCTTCTCTGGGCGCCGCGGCCAGCACA
TGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGACG
CAGCAGAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTG
GGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGC
TGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACACCGCCGTGCCCTGGAACGCCA
GCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC
GAGATCGACAACCTACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGA
GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACA
TCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGC
GCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCT
TCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGA
CGACCTGCGCAGCCTGTGCCTGTTAGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGC
CCGCATCGTGGAGCTGCTGGGCCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGC
TGCAGTACTGGATCCAGGAGCTGAAGAACAAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATC
GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCT
GCACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 26

SEQ ID NO:24 VAL120-THR202; ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCCTG
CCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCT
GAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGT
GCACCCACGGCATCCGCCCGGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAG
GAGGGCGTGGTGATCCGCAGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGCACT
GAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATACCA
TCGGCCCCCGCGCGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACT
GCAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCC
CAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGAT
GCACAGCTTCAACTGCGGCGGCGAGTTCCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC
CTGGAACAACACCATCGGCCCAACAACACCAACGGCACCATCACCTGCCCTGCCGATCA
AGCAGATCATCGGCGGCGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGC
AACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGAT
CTTCCGCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGG
TGGTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGACGCGC
GAGAAAGCGCGCCGTGACCCTGGGCGCCATGTTCTTGGGCTTCCTGGGCGCCGCGGCGCAGCACC
ATGGGCGCCCGCAGCCTGACCCTGACCCTGACCGTGACGGCCCGCAGCTGCTGAGCGGCATCGTGCA
GCAGCAGAAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGACGCTGACCGTGT
GGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAG
CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCC
AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCG
CGAGATCGACAACCTACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGG
AGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGAC
ATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGCGGCGCCTGGTGGGCCTG
CGCATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGC
TTCCAGACCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGG
CGGCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTGGCCCTGATCTGGG
ACGACCTGCGCAGCCTGTGCCTGTTACGCTACCAACCGCCTGCGCGACCTGATCCTGATCGCCG
CCCGCATCGTGAGCTGCTGGGCGCGCGGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG
CTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATCGCCATC
GCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCCAGCGCATCGGCCGCGCCTTCCT
GCACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 27

SEQ ID NO:25 VAL127-ASN195

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGTAACGGCGTGCCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTTGCAGCC
GGGGCAGGGAAGTCAACACCAGCGTGATCACCAGGCCTGCCCAAGGTGAGCTTCGAGCC
CATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT
CAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCACTGCACCCACGGCATCCGCCCCG
TGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC
GAGAACTTCACCGACAACGCCAAGACCATCATCGTGAGCTGAAGGAGAGCGTGGAGATCAA
CTGCACCCGCCCCAACAAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCTTCTA
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT
GGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGACCATC
GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG
CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCC
CAACAACACCAACGGCACCATCACCTGCCCTGCCGCATCAAGCAGATCATCAACCGCTGGC
AGGAGGTGGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAAC
ATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTT
CCGCCCCGGCGGCGGCGACATGCGCGCAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGG
TGAAGATCGAGCCCCCTGGGCGTGGCCCCCAAGGCCAAGCGCCGCGTGGTGACGCGCGAG
AAGCGCGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCCGCGGCAGCACCATG
GGCGCCCGCAGCCTGACCCTGACCGTGCAAGGCCCGCCAGCTGCTGAGCGGCATCGTGACGA
GCAGAACAACTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGTGGG
GCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTG
CTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAG
CTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCG
AGATCGACAACACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAG
AAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAAGTGGTTTCGACAT
CAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCG
CATCGTGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTT
CCAGACCCGCTTCCCCGCCCCCGCGGCCCGGACCGCCCCGAGGGCATCGAGGAGGAGGGCG
GCGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGAC
GACCTGCGCAGCCTGTGCCTGTTACGCTACCACCGCCTGCGCGACCTGATCCTGATCGCCGCC
CGCATCGTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCT
GCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATCGCCATCG
CCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCTTCCTGC
ACATCCCCCGCCGCATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAAGTTCGAG

FIG. 28

SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA
GTCTTCGTTTCGCCCCAGCGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG
TGGAAGGAGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCCAGGAGATCGTGCT
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG
GACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG
GGGGCAGGGAACTGCAACACCAGCGTGATACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCC
CATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTT
CAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAAGTGCAACCCACGGCATCCGCCCGG
TGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCAGC
GAGAACTTCACCGACAACGCCAAGACCATCATCGTGCAAGTGAAGGAGAGCGTGGAGATCAA
CTGCACCCGCCCCAACAAACACCCGCAAGAGCATCACCATCGGCCCGGCCGCGCCTTCTA
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAGAAGT
GGAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCCGGAACAAGACCATC
GTGTTCAAGCAGAGCAGCGGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG
CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCC
CAACAACACCAACGGCACCATCACCTGCCCTGCCGATCAAGCAGATCATCAACCGCGGCG
GCGGCAAGGCCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATCACC
GGCCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCCGCC
CGGGGGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAG
ATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGAGCGCGAGAAGCG
CGCCGTGACCCTGGGCGCCATGTTCTGGGCTTCTGGGCGCCGCCGCGCAGCACCATGGGCGC
CCGACGCTGACCCTGACCGTGACGGCCCGCCAGCTGCTGAGCGGCATCGTGAGCAGCAGCA
ACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGAGCTGACCGTGTGGGGCATC
AAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGG
CATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCAGCTGGA
GCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATC
GACAACTACCAACCTGATCTACACCCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAGAA
CGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCA
AGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTGGGCCTGCGCATCG
TGTTACCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGA
CCCGCTTCCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAG
CGCGACCGCGACCGCAGCAGCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGACCTG
CGCAGCCTGTGCCTGTTCAAGTACCAACCGCCTGCGCGACCTGATCCTGATCGCCGCCCGCATC
GTGGAGCTGCTGGGCCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCAGTA
CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCCATCGCCATCGCCGTGG
CCGAGGGCACCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCGCCTTCTGCACATCC
CCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACCTCGAG

FIG. 29

SEQUENCE LISTING

<110> Chiron Corporation

<120> MODIFIED HIV ENV POLYPEPTIDES

<130> 1605.100

<140>

<141>

<160> 26

<170> PatentIn Ver. 2.0

<210> 1

<211> 856

<212> PRT

<213> Human immunodeficiency virus

<400> 1

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg
 1 5 10 15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu
 20 25 30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 35 40 45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
 50 55 60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
 85 90 95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
 100 105 110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser
 115 120 125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser
 130 135 140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn
 145 150 155 160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe
 165 170 175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys
 180 185 190

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val
 195 200 205
 Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala
 210 215 220
 Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr
 225 230 235 240
 Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser
 245 250 255
 Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile
 260 265 270
 Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu
 275 280 285
 Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg
 290 295 300
 Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile
 305 310 315 320
 Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala
 325 330 335
 Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln
 340 345 350
 Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp
 355 360 365
 Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr
 370 375 380
 Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp
 385 390 395 400
 Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu
 405 410 415
 Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys
 420 425 430
 Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn
 435 440 445
 Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu
 450 455 460
 Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
 465 470 475 480
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val
 485 490 495
 Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala
 500 505 510

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser
 515 520 525
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu
 530 535 540
 Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu
 545 550 555 560
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu
 565 570 575
 Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu
 580 585 590
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val
 595 600 605
 Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn
 610 615 620
 His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser
 625 630 635 640
 Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn
 645 650 655
 Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp
 660 665 670
 Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile
 675 680 685
 Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile
 690 695 700
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His
 705 710 715 720
 Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu
 725 730 735
 Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser
 740 745 750
 Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr
 755 760 765
 His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu
 770 775 780
 Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu
 785 790 795 800
 Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn
 805 810 815
 Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val
 820 825 830

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg
 835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu
 850 855

<210> 2

<211> 847

<212> PRT

<213> Human immunodeficiency virus

<400> 2

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Gly
 1 5 10 15

Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys
 20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
 35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
 65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
 85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
 100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
 115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp
 130 135 140

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr
 145 150 155 160

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys
 165 170 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile
 180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe
 195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu
 210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val
 225 230 235 240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln
 245 250 255
 Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser
 260 265 270
 Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu
 275 280 285
 Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser
 290 295 300
 Ile Thr Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile
 305 310 315 320
 Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn
 325 330 335
 Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn
 340 345 350
 Lys Thr Ile Val Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val
 355 360 365
 Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr
 370 375 380
 Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr
 385 390 395 400
 Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg
 405 410 415
 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln
 420 425 430
 Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
 435 440 445
 Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly
 450 455 460
 Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val
 465 470 475 480
 Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val
 485 490 495
 Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly
 500 505 510
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu
 515 520 525
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn
 530 535 540
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr
 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg
 565 570 575
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys
 580 585 590
 Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys
 595 600 605
 Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg
 610 615 620
 Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser
 625 630 635 640
 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys
 645 650 655
 Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr
 660 665 670
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile
 675 680 685
 Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser
 690 695 700
 Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg
 705 710 715 720
 Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser
 725 730 735
 Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg
 740 745 750
 Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile
 755 760 765
 Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu
 770 775 780
 Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn
 785 790 795 800
 Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly
 805 810 815
 Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu
 820 825 830
 His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu
 835 840 845

<210> 3

<211> 2310

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Ala204

<400> 3

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcgcc 360
ggcgcctgcc ccaaggtgag ctctgagccc atccccatcc actactgcgc ccccgccggc 420
ttcgccatcc tgaagtgcaa cgacaagaag ttcaacggca gcggcccctg caccaacgtg 480
agcaccgtgc agtgacacca cggcatccgc cccgtggtga gcaccagct gctgtgaac 540
ggcagcctgg ccgaggaggg cgtggtgatc cgcagcgaga acttcaccga caacgccaag 600
accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcaccgccc caacaacaac 660
acccgcaaga gcatcaccat cggccccggc cgcgccttct acgccaccgg cgacatcatc 720
ggcgacatcc gccaggccca ctgcaacatc agcggcgaga agtggaaaca caccctgaag 780
cagatcgtga ccaagctgca ggcccagttc ggcaacaaga ccacgtgtt caagcagagc 840
agcggcgggc accccgagat cgtgatgcac agcttcaact gcggcgggca gttcttctac 900
tgcaacagca cccagctgtt caacagcacc tggaaacaaca ccacggccc caacaacacc 960
aacggcacca tcaccctgcc ctgccgcac aagcagatca tcaaccgctg gcaggaggtg 1020
ggcaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080
ggcctgctgc tgaccgcga cggcggaag gagatcagca acaccaccga gatcttccgc 1140
cccgcgggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200
aagatcgagc ccctgggcgt ggccccacc aaggccaagc gccgcgtggt gcagcgcgag 1260
aagcgcgccg tgaccctggg cgccatgttc ctgggcttcc tggcgccgc cggcagcacc 1320
atgggcgccc gcagcctgac cctgaccgtg caggcccgcc agctgctgag cggcatcgtg 1380
cagcagcaga acaacctgct gcgcgccatc gaggccagc agcacctgct gcagctgacc 1440
gtgtggggca tcaagcagct gcaggcccgc gtgctggccg tggagcgcta cctgaaggac 1500
cagcagctgc tgggcatctg gggctgcagc ggcaagctga tctgcaccac cgccgtgccc 1560
tggaaacgcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gcgagatcga caactacacc aacctgatct acaccctgat cgaggagagc 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaaactggt tcgacatcag caagtggctg tggatcatca agatcttcat catgatcgtg 1800
ggcggcctgg tgggcctgcg catcgtgttc accgtgctga gcacgtgaa ccgcgtgcgc 1860
cagggctaca gccccctgag cttccagacc cgcttccccg ccccccggg ccccgaccgc 1920
cccgagggca tcgaggagga gggcggcgag cgcgaccgcg accgcagcag cccctggtg 1980
cacggcctgc tggccctgat ctgggacgac ctgcgcagcc tgtgcctgtt cagctaccac 2040
cgccctgcgc acctgatcct gatcgccgcc cgcatcgtgg agctgctggg ccgcgcggc 2100
tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agcgccgtga gcctgttcga cgccatcgcc atcgccgtgg ccgagggcac cgaccgcac 2220
atcgaggtgg cccagcgcac cggccgcgcc ttccctgcaca tcccccgccg catccgccag 2280
ggcttcgagc gcgcctgct gtaactcgag 2310

```

<210> 4

<211> 2316

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Ile201

<400> 4

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360

```

```

atcaccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatacctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtagagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatacgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcataaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcgcc agatccgctg cagcagcaac 1080
atcaccggcc tgctgctgac ccgcgacggc ggcaaggaga tcagcaaac caccgagatc 1140
ttccgccccg tggcgccgga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagccctt gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctggcgccc atgttctctg gcttctctgg cgccgcccgc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggccc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcata cagaccgct tccccgcccc ccgcgcccc 1920
gtgcgccagg gctacagccc cctgagcttc cagaccgct tccccgcccc ccgcgcccc 1980
gaccgccccg agggcatcga ggaggaggcc ggcgagcgcg accgcgaccg cagcagcccc 2040
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2100
taccaccgcc tgcgcgacct gatcctgac gccgcccgc tctgaggagt gctggggcgc 2160
cgcggtggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2220
aagaacagcg ccgtgagcct gttegacgcc atcgccatcg ccgtggccga gggcaccgac 2280
cgcatcatcg aggtggccca gcgcacggc cgcgccttcc tgcacatccc ccgcccgcac 2316
cgccagggct tcgagcgcgc cctgctgtaa ctcgag

```

<210> 5

<211> 2322

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Ile201B

<400> 5

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggccctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcaccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatacctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtagagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatacgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960

```

```

aacaccaacg gcaccatcac cctgccctgc cgcacaaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cggccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcaccggcc tgetgctgac ccgcgacggc ggcaaggaga tcagcaaacac caccgagatc 1140
ttccgccccg gcggcgcgca catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagccccct gggcggtggc cccaccaagg ccaagcgccg cgtggtgacg 1260
cgcgagaagc gcgccgtgac cctggggcgcc atgttccctgg gcttccctggg cgcgcggcgc 1320
agcaccatgg gcgcccgcag cctgacctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgctg tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1560
gtgccctgga acgccagctg gagcaacaag agcctggacc agatctgga caacatgacc 1620
tggatggagt gggagcgcgca gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggccc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtgggcg gcctggtggg cctgcgcacg gtgttcaccg tgcctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagaccgcgt tccccgcccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg acccgagccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccga tcgtggagct gctgggcccgc 2100
cgcggtggtg aggccctgaa gtactggggc aacctgctgc agtactgat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcacggc cgcgccttcc tgcacatccc ccgcgcgcatc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgagcgtg ct 2322

```

<210> 6

<211> 2328

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Lys121-Val200

<400> 6

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttccgcccag cgcgctggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggccctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgctg tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
cccgatgatc cccaggcctg cccaagggtg agcttcgagc ccaccccat cactactgac 420
gccccgcgcg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcgcccc 480
tgcaccaacg tgagcacctg gcagtgcacc cacggcatcc gcccgtggt gagcaccag 540
ctgctgctga acggcagcct ggccgaggag ggctggtgta tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcaccgc 660
cccaacaaca acaccgcaa gagcatcacc atcgccccg gccgcgctt ctacgccacc 720
ggcgacatca tcggcgacat ccgccaggcc cactgcaaca tcagcggcga gaagtgaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaagcaga gcagcggcg cgaccccag atcgtgatgc acagcttcaa ctgcggcggc 900
gagttcttct actgcaacag caccagctg ttcaacagca cctggaacaa caccatcggc 960
cccaacaaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaaccgc 1020
tggcaggagg tgggcaaggc catgtacgcc ccccccaccc gcggccagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgaccgcg gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gccccggcg cggcgacatg cgcgacaact ggcgagcga gctgtacaag 1200
tacaaggtgg tgaagatcga gccctggggc gtggcccca ccaaggccaa gcgcgcgctg 1260
gtgcagcgcg agaagcgcg cgtagccctg ggcgccatgt tcctgggctt cctgggcgccc 1320
gccggcagca ccatgggcgc ccgcagcctg accctgaccg tgcaggcccc ccagctgctg 1380
agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcac tggggctgca gcggcaagct gatctgcacc 1560

```

```

accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcaggagaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcacgtgtg tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgaggg catcgaggag gagggcgggc agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcacgtg ggagctgctg 2100
ggccgcccgc gctgggaggg cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgccatcg ccatcgccgt ggccgagggc 2220
accgaccgca tcatcgaggt ggcccagcgc atcgcccgcg ccttcctgca catccccccg 2280
cgcacccgcc agggcttcga gcgcgccctg ctgtaactcg agcgtgct 2328

```

<210> 7

<211> 2334

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Leu122-Ser199

<400> 7

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgtgga aggaggccac caccacctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggccctgccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgcccggctt cgccatcctg aagtgcacg acaagaagt caacggcagc 480
ggcccctgca ccaacgtgag caccgtgcag tgcacccacg gcacccccc cgtggtgagc 540
accagctgc tgctgaacgg cagcctggcc gagggggcg tggtgatccg cagcgagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
acccgcccc acaacaacac ccgcaagagc atcaccatcg gccccggccg cgcttcttac 720
gccaccggcg acatcatcgg cgacatccgc caggccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgggcag cccgagatcg tgatgcacag cttcaactgc 900
ggcgggcaggt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcgccccca acaacaccaa cggcaccatc accctgccct gccgcacaa gcagatcatc 1020
aaccgctggc aggaggtggg caaggccatg tacgcccccc ccatccgcgg ccagatccgc 1080
tgcagcagca acatcacccg cctgctgctg acccgcgacg gcggcaagga gatcagcaac 1140
accaccgaga tcttcgccc cggcgggcgg gacatgcgcg acaactggcg cagcgagctg 1200
tacaagtaca aggtggtgaa gatcgagccc ctggggcgtg cccccaccaa ggccaagcgc 1260
cgcgtggtgc agcgcgagaa gcgcgccgtg accctggcg ccatgttcct gggcttctctg 1320
ggcgccgccc gcagcaccat gggcgcccgc agcctgaccc tgaccgtgca ggcccgcag 1380
ctgctgagcg gcacgtgca gcagcagaac aacctgctgc gcgccatcga ggcccagcag 1440
cacctgctgc agctgaccgt gtggggcatc aagcagctgc agggcccgct gctggccgtg 1500
gagcgctacc tgaaggacca gcagctgctg ggcacatggg gctgcagcgg caagctgac 1560
tgcaccaccg ccgtgccctg gaacgccagc tggagcaaca agagcctgga ccagatctgg 1620
aacaacatga cctggatgga gtgggagcgc gagatcgaca actacaccaa cctgatctac 1680
accctgatcg aggagagcca gaaccagcag gagaagaacg agcaggagct gctggagctg 1740
gacaagtggg ccagcctgtg gaactggttc gacatcagca agtggctgtg gtacatcaag 1800
atcttcatca tgatcgtggg cggcctggtg ggcctgcgca tcgtgttcac cgtgctgagc 1860
atcgtgaacc gcgtgcgcca gggctacagc cccctgagct tccagaccgg cttccccgcc 1920
cccccgggcc ccgaccgccc cgagggcacg gagggaggag gcggcgagcg cgaccgcgac 1980
cgcagcagcc ccctggtgca cggcctgctg gcctgatct gggacgacct gcgcagcctg 2040
tgccctgttca gctaccaccg cctgcgcgac ctgatcctga tcgcgcgccg catcgtggag 2100
ctgctggggc gcccgggctg ggaggccctg aagtactggg gcaacctgct gcagtactgg 2160

```

```

atccaggagc tgaagaacag cgccgtgagc ctgttcgacg ccatcgccat cgccgtggcc 2220
gagggcaccg accgcatcat cgaggtggcc cagcgcacg gccgcgctt cctgcacatc 2280
ccccgccgca tccgccaggg cttcgagcgc gccctgctgt aactcgagcg tgct 2334

```

<210> 8

<211> 2316

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Thr202

<400> 8

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
gccaccagg cctgcccacaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcggt gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
catgacgagc gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccga gctgttcaac agcaactgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcacaaagc agatcatcaa ccgctggcag 1020
gaggtgggca aggccatgta cgccccccc atccgcggcc agatccgctg cagcagcaac 1080
atcacgggca tgctgctgac ccgcgacggc ggcaaggaga tcagcaacac caccgagatc 1140
ttccgccccg gcggcgcgca catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaga tcgagcccct gggcgtaggc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgccgtgac cctgggcgcc atgttcctgg gcttcctggg cgccgcccgc 1320
agcaccatgg gcgcccgcag cctgaccctg accgtgcagg cccgccagct gctgagcggc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgaccgtgt ggggcatcaa gcagctgcag gcccgcgtgc tggccgtgga gcgtacctg 1500
aaggaccagc agctgctggg catctggggc tgcagcgcca agctgatctg caccaccgcc 1560
gtgccttggg acgcccagct gagcaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgga gatcgacaac tacaccaacc tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggca 1740
agcctgtgga actggttcga catcagcaag tggctgtggg acatcaagat cttcatcatg 1800
atcgtgggag gcctggtggg cctgcgcac gtgttcaccg tgctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagaccgct tccccgccc ccgcggcccc 1920
gaccgccccg agggcatcga ggaggagggc ggcgagcgcg accgcgaccg cagcagcccc 1980
ctggtgcacg gcctgctggc cctgatctgg gacgacctgc gcagcctgtg cctgttcagc 2040
taccaccgcc tgcgcgacct gatcctgatc gccgcccga tcgtggagct gctgggcccgc 2100
cgcggtggg aggccctgaa gtactggggc aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atcgccatcg ccgtggccga gggcaccgac 2220
cgcatcatcg aggtggccca gcgcacggc cgcgccttc tgacatccc ccgcccgcac 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag 2316

```

<210> 9

<211> 2541

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Trp427-Gly431

<400> 9

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttegccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgctgct ccaccgacce caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200
ggcaccatca ccctgccttg ccgcataaag cagatcatca accgctgggg cggaaggcc 1260
atgtacgccc ccccatcgg cgccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgaccgcg acggcggaac ggagatcagc aacaccaccg agatcttcg ccccgccggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaagtggtg gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgca gaagcgcgcc 1500
gtgaccctgg gcgccatgtt cctgggcttc ctgggcgccc ccggcagcac catgggcgccc 1560
cgcagcctga ccctgaccgt gcaggcccg cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgac tacacctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcgccctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccaggggctac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gacctgatcc tgatcgccgc ccgcacgtg gagctgctgg gccgcgcgg ctgggaggcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccctg 2400
agcctgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgcat catcgagggt 2460
gcccagcgca tcggccgcgc ctccctgcac atccccgcc gcatccgcca gggcttcgag 2520
cgcgccctgc tgtaactcga g

```

<210> 10

<211> 2541

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Arg426-Gly431

<400> 10

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttegccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180

```



```

accgaggtgc acaacgtgtg ggccacccac gcctgctgtc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcaactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gectgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggttct 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caactgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg agggggcggt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tctgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccg agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcgg cggcaaggcc 1260
atgtacgccc ccccatccg cgccagatc cgtgcagca gcaacatcac cggctgctg 1320
ctgaccgcg acggcggaag ggagatcagc aacaccaccg agatcttccg ccccgccggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaagggtgt gaagatcgag 1440
ccccggggcg tggccccac caaggccaag cgccgctgtg tgcagcgca gaagcgcgcc 1500
gtgaccctgg gcgcatgtt cctgggcttc ctgggcgcg ccggcagcac catgggcgcc 1560
cgcagcctga ccctgaccgt gcaggccgc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgcat cgaggcccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggccccg cgtgctggcc gtggagcgt acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgctgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacacctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcgccctg 2040
gtgggcctgc gcatcgtgtt caccgtgctg agcatcgtga accgcgtgcg ccagggttac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcggcga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgctgcgc 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgcgcggt ctggagggcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcctgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
gccccagcga tcggccgcgc ctctctgcac atccccgcc gcatccgcca gggcttcgag 2520
cgcgccctgc tgtaactcga g

```

<210> 11

<211> 2541

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Arg426-Gly431B

<400> 11

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgctgtc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcaactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540

```

```

atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgctg agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gccttctacg ccaccggcga catcatcgcc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcggcag cggcaaggcc 1260
atgtacgccc ccccatccg cgccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcggaac ggagatcagc aacaccaccg agatcttccg ccccgcgcg 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaagtggtg gaagatcgag 1440
cccctgggcg tggccccac caaggccaag cgccgcgtgg tgcagcgcg gaagcgcgcc 1500
gtgacctgg gcgcatgtt cctgggcttc ctgggcgccc ccggcagcac catggcgccc 1560
cgcagcctga ccctgaccgt gcaggcccg cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgccgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacacctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcgccctg 2040
gtgggctgct gcatcggtt caccgtgctg agcatcgtga accgcgtgcg ccagggtac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgaccg ccccgagggc 2160
atcgaggagg agggcgcgca gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgctgcg 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgcccgg ctgggaggcc 2340
ctgaagtact ggggcaacct gctgcagtag tggatccagg agctgaagaa cagcgccgtg 2400
agcctgttcg acgcatcgc catcgccgtg gccgagggca ccgaccgcat catcgaggtg 2460
gcccagcgca tcggccgcgc cttcctgcac atcccccgcc gcatccgcca gggcttcgag 2520
cgcgccctgc tgtaactcga g

```

<210> 12

<211> 2541

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Arg426-Lys432

<400> 12

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccggtgtgga aggaggccac caccacctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgctg agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900

```

```

cgcaagagca tcaccatcgg ccccgccgcg gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgccgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc cagctgttcaa cagcacctgg aacaacacca tcggcccca caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca accgcgccgg caacaaggcc 1260
atgtacgccc ccccatccg cgccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgacccgcg acggcgccaa ggagatcagc aacaccaccg agatcttcg ccccgccggc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaagggtgt gaagatcgag 1440
cccctgggcg tggcccccac caaggccaag cgccgcgtgg tgcagcgga gaagcgccg 1500
gtgaccctgg gcgccatgtt cctgggcttc ctggcgccg ccggcagcac catggcgcc 1560
cgcagcctga ccctgaccgt gcaggcccg cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgcat cgaggccag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgctgtgcc ctggaacgcc 1800
agctggagca acaagagcct ggaccagatc tggacaaca tgacctggat ggagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacacctga tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactgg 1980
ttcgacatca gcaagtggct gtggtacatc aagatcttca tcatgatcgt gggcgccctg 2040
gtgggctgct gcacgtgtt caccgtgctg agcatcgtga accgctgctg ccagggtac 2100
agccccctga gcttccagac ccgcttcccc gccccccg cgccccgacc ccccgagggc 2160
atcgaggagg agggcgccga gcgcgaccgc gaccgcagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgctgcgc 2280
gacctgatcc tgatcgccgc ccgcacgtg gagctgctgg gccgcgcgg ctgggaggcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcctgttcg acgccatcgc catcgccgtg gccgagggca ccgaccgat catcgaggtg 2460
gcccgagcga tcggccgcgc ctccctgcac atccccgc gcacccgcca gggcttcgag 2520
cgcgcctgcg tgaactcga g

```

<210> 13

<211> 2535

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Asn425-Lys432

<400> 13

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
ccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgctgct ccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gactgcacc aacctgaaga acgccacca caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagct gatcaccag 600
gcctgcccc aagtgagctt cgagcccatc cccatccact actgcgccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca ccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgaca gcccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgccgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc cagctgttcaa cagcacctgg aacaacacca tcggcccca caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcatca acgccccaa ggccatgtac 1260

```

```

cccccccca tccgcggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgcccccg cggcgggcag 1380
atgcgcgaca actggcgag cgagctgtac aagtacaagg tggtagaat cgagccccctg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgcctgacc 1500
ctgggcgcca tgttcctggg cttcctgggc gccgcccggc gcaccatggg cgcctgcagc 1560
ctgaccctga ccgtgcaggc ccgccagctg ctgagcggca tcgtgcagca gcagaacaac 1620
ctgctgcgcg ccacgcaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1680
cagctgcagg cccgcgtgct ggccgtggag cgctacctga aggaccagca gctgctgggc 1740
atctggggct gcagcggcaa gctgatctgc accaccgccc tgccctggaa cgcagctgg 1800
agcaacaaga gcctggacca gatctggaac aacatgacct ggatggagt ggagcgcgag 1860
atcgacaact acaccaacct gatctacacc ctgatcgagg agagccagaa ccagcaggag 1920
aagaacgagc aggagctgct ggagctggac aagtgggcca gcctgtggaa ctgggttcgac 1980
atcagcaagt ggctgtggtg catcaagatc ttcacatga tcgtggggcg cctgggtggc 2040
ctgcgcctcg tgttcaccgt gctgagcatc gtgaaccgcg tgcgccaggg ctacagcccc 2100
ctgagcttcc agaccgctt cccgcccccg cgcggccccg accgccccga gggcatcgag 2160
gaggaggggc gcgagcgcga ccgcgaccgc agcagcccc tggtgcacgg cctgctggcc 2220
ctgatctggg acgacctgcg cagcctgtgc ctgttcagct accaccgct gcgcgacctg 2280
atcctgatcg cgcctgcag ctgtggagctg ctgggcccgc gcgctggga ggccctgaag 2340
tactggggca acctgctgca gtactggatc caggagctga agaacagcgc cgtgagcctg 2400
ttcgacgcca tcgccatcgc cgtggccgag ggcaccgacc gcatcatcga ggtggcccag 2460
cgcacggccc gcgccttct gcacatcccc cgcgcacatc gccagggctt cgagcgcgccc 2520
ctgctgtaac tcgag 2535

```

<210> 14

<211> 2529

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ile424-Ala433

<400> 14

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgcctggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgtgtgga aggaggccac caccacctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgctgct ccaccgacc caacccccag 240
gagatcgctg tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgggtggg 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagcctg cgtgaagctg 360
acccccctgt gcgtgacct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtag gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagct gatcaccag 600
gcctgcccc aagtgagctt cgagcccatc cccatccact actgcgccc cgcggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggaggcggt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgctg agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgggcagc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200
ggcaccatca ccctgcccgt ccgcatcaag cagatcatcg gcggcgccat gtacgcccc 1260
cccacccgag gccagatccg ctgcagcagc aacatcacgg gcctgctgct gaccgcgac 1320
ggcggaagg agatcagcaa caccaccgag atcttcgccc ccggcgggcg cgacatgcgc 1380
gacaactggc gcagcgagct gtacaagtag aaggtggtga agatcgagcc cctgggcgtg 1440
gccccacca aggccaagcg ccgcgtggtg cagcgcgaga agcgcgccgt gaccctgggc 1500
gccatgttcc tgggcttctt gggcgccgccc ggcagacca tggcgcccc cagcctgacc 1560
ctgaccgtgc agggccgcca gctgctgagc ggcacgtgct agcagcagaa caacctgctg 1620

```

```

cgcgccatcg agggccagca gcacctgctg cagctgaccg tgtggggcat caagcagctg 1680
cagggcccgcg tgctggccgt ggagcgctac ctgaaggacc agcagctgct gggcatctgg 1740
ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacgccag ctggagcaac 1800
aagagcctgg accagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac 1860
aactacacca acctgatcta caccctgatc gaggagagcc agaaccagca ggagaagaac 1920
gagcaggagc tgctggagct ggacaagtgg gccagcctgt ggaactgggt cgacatcagc 1980
aagtggctgt ggtacatcaa gatcttcatc atgatcgtgg gcggcctggt gggcctgcgc 2040
atcgtgttca ccgtgctgag catcgtgaac cgcgtgcgcc agggctacag cccctgagc 2100
ttccagaccc gcttccccgc ccccgcgggc cccgaccgcc ccgaggcat cgaggaggag 2160
ggcggcgagc gcgaccgca ccgcagcagc cccctggtgc acggcctgct ggcctgatc 2220
tgggacgacc tgcgcagcct gtgcctgttc cgtaccacc gcctgcgcga cctgatcctg 2280
atcgccgccc gcacgtgga gctgctgggc cgccgcggt gggaggccct gaagtactgg 2340
ggcaaacctgc tgcagtactg gatccaggag ctgaagaaca gcgcgtgag cctgttcgac 2400
ggcatcgcca tcgccgtggc cgagggcacc gaccgcatca tcgaggtggc ccagcgcac 2460
ggccgcgcct tcctgcacat ccccgcgcg atccgccagg gcttcgagcg cgccctgctg 2520
taactcgag 2529

```

<210> 15

<211> 2523

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Ile423-Met434

<400> 15

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttccgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgacce caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgccaccaa caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggaggcggt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca ccctgccctg ccgcatcaag cagatcgggc gcatgtacgc ccccccatc 1260
cgcgggccaga tccgctgcag cagcaacatc accggcctgc tgctgaccgg cgacggcggc 1320
aaggagatca gcaacaccac cgagatcttc cgccccggcg gcggcgacat gcgcgacaac 1380
tggcgcgagc agctgtacaa gtacaagggt gtgaagatcg agccctggg cgtggccccc 1440
accaaggcca agcgcgcgt ggtgcagcgc gagaagcgcg ccgtgacct gggcgccatg 1500
ttcctgggct tcctgggcgc cgccggcagc acctggggcg cccgcagcct gaccctgacc 1560
gtgcaggccc gccagctgct gagcggcatc gtgcagcagc agaacaacct gctgcgcgcc 1620
atcgaggccc agcagcacct gctgcagctg accgtgtggg gcatcaagca gctgcaggcc 1680
cgcggtgctg ccgtggagcg ctacctgaag gaccagcagc tgctgggcat ctggggctgc 1740
agcggcaagc tgatctgcac caccgccgtg ccctggaacg ccagctggag caacaagagc 1800
ctggaccaga tctggaacaa catgacctgg atggagtggg agcgcgagat cgacaactac 1860
accaacctga tctacacct gatcgaggag agccagaacc agcaggagaa gaacgagcag 1920
gagctgctgg agctggacaa gtgggcccag ctgtggaact ggttcgacat cagcaagtgg 1980

```

```

ctgtggtaca tcaagatctt catcatgac gtgggcggcc tgggtgggct ggcgcatcgtg 2040
ttcacccgtgc tgagcatcgt gaaccgcgtg cgccagggtc acagccccct gagcttccag 2100
acccgcttcc ccgccccccg cggccccgac cgccccgagg gcatcgagga ggagggcggc 2160
gagcgcgacc gcgaccgcag cagccccctg gtgcacggcc tgctggccct gatctgggac 2220
gacctgcgca gcctgtgcct gttagctac caccgcctgc gcgacctgat cctgatcgcc 2280
gcccgcacgc tggagctgct gggccgcccg ggctgggagg ccctgaagta ctggggcaac 2340
ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgacgccatc 2400
gccatcgccg tggccgaggg caccgaccgc atcatcgagg tggcccagcg catcgccgc 2460
gccttccctgc acatcccccg ccgcatccgc cagggcttcg agcgcgccct gctgtaactc 2520
gag 2523

```

<210> 16

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Gln422-Tyr435

<400> 16

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttccgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggccctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgctgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgccacca caccaagagc 420
agcaactcga aggagatgga ccgcgccgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtag gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgccccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgctgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gccttctacg ccaccggcga catcatcggc 960
gacatccgcc aggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgcgacc ccgagatcgt gatgcacagc ttcaactcgc gcggcgagtt cttctactgc 1140
aacagcacc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca cctgcctcgc ccgcatcaag cagggcggtc acgccccccc catccgcggc 1260
cagatccgct gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcgggcgcg acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tggcggtggc ccccaaccaag 1440
gccaagcgcc gcgtggtgca gcgcgagaag cgcgccgtga ccctgggcgc catgttcctg 1500
ggcttccctgg gcgcgcggcg cagcaccatg ggcgcccgca gcctgaccct gaccgtgcag 1560
gcccgcacgc tgctgagcgg catcgtgcag cagcagaaca acctgctgcg cgccatcgag 1620
gcccagcagc acctgctgca gctgaccgtg tggggcatca agcagctgca ggcccgcgtg 1680
ctggccgtgg agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccaccgc cgtgcctcgg aacgccagct ggagcaacaa gagcctggac 1800
cagatctgga acaacatgac ctggatggag tgggagcgcg agatcgacaa ctacaccaac 1860
ctgatctaca ccctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggctgtgg 1980
tacatcaaga tcttcatcat gatcgtgggc ggcctgggtg gcctgcgcac cgtgttcacc 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagaccgcg 2100
ttccccgccc cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggctgctgg ccctgatctg ggacgacctg 2220
cgagcctgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atcgtggagc tgctgggccc ccgcggtgg gagggcctga agtactgggg caacttgctg 2340

```

```

cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
gccgtggccg agggcaccga ccgcatcatc gaggtggccc agcgcatcgg ccgcgccttc 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag 2517

```

<210> 17

<211> 2517

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Gln422-Tyr435B

<400> 17

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgaccct gactgcacc aacctgaaga acgccacca caccaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccag 600
gcctgcccc aagtgagctt cgagcccatc cccatccact actgcgccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gcccctgcac caacgtgagc 720
accgtgcagt gcaccacgg catccgccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggaggcggt ggtgatccg agcgagaact tcaccgaaa cgccaagacc 840
atcactgtgc agctgaagga gagctggag atcaactgca cccgcccac caacaacacc 900
cgcaagagca tcaccatcgg ccccgggcgc gccttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgggcagc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt cttctactgc 1140
aacagcacc agctgttcaa cagcacctgg aacaacacca tcggcccaa caacaccaac 1200
ggcaccatca cctgccctg ccgcatcaag caggccccct acgcccccc catccgcggc 1260
cagatccgct gcagcagcaa catcaccggc ctgctgctga cccgcgacgg cggcaaggag 1320
atcagcaaca ccaccgagat cttccgcccc ggcggcgggc acatgcgcga caactggcgc 1380
agcgagctgt acaagtacaa ggtggtgaag atcgagcccc tgggcgtggc cccaccaag 1440
gccaagcgcc gcgtggtgca gcgcgagaag cgcgccgtga ccctgggcgc catgttctctg 1500
ggcttctctg gcgcggccgg cagcaccatg ggcgcccga gcctgacct gaccgtgcag 1560
gcccgcagc tgctgagcgg catcgtgcag cagcagaaca acctgctgcg cgccatcgag 1620
ctggcctgag agcgctacct gaaggaccag cagctgctgg gcatctgggg ctgcagcggc 1740
aagctgatct gcaccaccgc cgtgccctgg aacgccagct ggagcaacaa gagcctggac 1800
cagatctgga acaacatgac ctggatggag tgggagcgcg agatcgaaa ctacaccaac 1860
ctgatctaca cctgatcga ggagagccag aaccagcagg agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggttcg acatcagcaa gtggtgtggtg 1980
tacatcaaga tttcatcat gatcgtgggc ggcctggtgg gcctgcgcac cgtgttcacc 2040
gtgctgagca tcgtgaaccg cgtgcgccag ggctacagcc ccctgagctt ccagaccgc 2100
ttccccgcc cccgcggccc cgaccgccc gagggcatcg aggaggaggg cggcgagcgc 2160
gaccgcgacc gcagcagccc cctggtgcac ggctgctgg ccctgatctg ggacgacctg 2220
cgcagcctgt gcctgttcag ctaccaccgc ctgcgcgacc tgatcctgat cgccgcccgc 2280
atcgtggagc tgctgggccc ccgcggctgg gagggcctga agtactgggg caacctgctg 2340
cagtactgga tccaggagct gaagaacagc gccgtgagcc tgttcgacgc catcgccatc 2400
gccgtggccg agggcaccga ccgcatcatc gaggtggccc agcgcatcgg ccgcgccttc 2460
ctgcacatcc cccgccgcat ccgccagggc ttcgagcgcg ccctgctgta actcgag 2517

```

<210> 18

<211> 2322

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Leu122-Ser199;
Arg426-Gly431

<400> 18

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacce caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggcctgcccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgcccggctt cgccatcctg aagtgcacag acaagaagtt caacggcagc 480
ggccccctgca ccaacgtgag caccgtgcag tgcaccacag gcatccgccc cgtggtgagc 540
accagctgc tgcgaacgg cagcctggcc gaggagggcg tggatgaccg cagcgagAAC 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
acccgcccc acaacaacac ccgcaagagc atcaccatcg gccccggccg cgccttctac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tgaacaaca ccctgaagca gatcgtgacc aagctgcagg ccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgggcag cccgagatcg tgatgcacag cttcaactgc 900
ggcgggcagat tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcgggccca acaacaccaa cggcaccatc accctgccct gccgcatcaa gcagatcatc 1020
aaccgcgcg cgggcaaggc catgtacgcc cccccatcc cgggccagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgaccgc gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gccccggcg cgcgacatg cgcgacaact ggcgacgca gctgtacaag 1200
tacaagttgg tgaagatcga gcccctgggc gtggcccca ccaaggcca gcgcgcgtg 1260
gtgcagcgcg agaagcgcg cgtgaccctg ggcgccatgt tcctgggctt cctgggcgc 1320
gccggcagca ccatgggcgc ccgcagcctg accctgaccg tgcaggcccg ccagctgctg 1380
agcggcacatg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcac tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacacctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcacgtgtg tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cccccccgc 1920
ggccccgacc gccccgagg catcgaggag gaggcgcgcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg cccgcacatg ggagctgctg 2100
ggccgcccgc gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgccatcg ccacgcccgt ggccgagggc 2220
accgaccgca tcacgaggt ggcccagcgc atcgccgcg ccttcctgca catccccgc 2280
cgcacccgcc agggcttcga gcgcgccctg ctgtaactcg ag 2322

```

<210> 19

<211> 2322

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Leu122-Ser199;
Arg426-Lys432

<400> 19

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120

```



```

cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggccctgccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgaacg acaagaagtt caacggcagc 480
ggccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
accagctgc tgctgaacgg cagcctggcc gaggagggcg tggatccg cagcgagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accgcacca acaacaacac ccgcaagagc atcaccatcg gccccggccg cgccttctac 720
gccaccggcg acatcatcgg cgacatccgc caggccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg ccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcggcgac ccgagatcg tgatgcacag cttcaactgc 900
ggcggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcggcccca acaacaccaa cggcaccatc accctgccct gccgcatcaa gcgatcatc 1020
aaccgcggcg gcaacaaggc catgtacgcc cccccatcc gccggcagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgaccgc gacggcgga aggagatcag caacaccacc 1140
gagatcttcc gccccggcgg cggcgacatg cgcgacaact ggcgagcga gctgtacaag 1200
tacaaggtgg tgaagatcga gcccctgggc gtggcccca ccaaggccaa gcgccgcgtg 1260
gtgcagcgcg agaagcgcg cgtgaccctg ggcgccatgt tcctgggctt cctgggcgcc 1320
gccggcagca ccatgggcgc ccgcagcctg accctgaccg tcgaggcccg ccagctgctg 1380
agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcacg tggggctgca gcggcaagct gatctgcacc 1560
accgcctgct cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaa 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tgggcggcct ggtgggcctg cgcacgtgtg tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttcaga cccgcttccc cgcccccg 1920
ggccccgacc gccccgaggg catcgaggag gaggggcgcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgate ctgatcgccg ccgcacatcg ggagctgctg 2100
ggccgcccgc gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcotgttc gacgccatcg ccacgcccgt ggccgagggc 2220
accgaccgca tcacgaggt ggcccagcgc atcgggccg ccttcctgca catccccgc 2280
cgcacccgcc agggcttcga gcgcgccctg ctgtaactcg ag 2322

```

<210> 20

<211> 2322

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Leu122-Ser199;
Trp427-Gly431

<400> 20

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggccctgccc aaggtgagct tcgagcccat ccccatccac 420
tactgcgccc ccgccggctt cgccatcctg aagtgaacg acaagaagtt caacggcagc 480
ggccctgca ccaacgtgag caccgtgcag tgcacccacg gcatccgccc cgtggtgagc 540
accagctgc tgctgaacgg cagcctggcc gaggagggcg tggatccg cagcgagaac 600
ttcaccgaca acgccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660

```

```

acccgcccc acaacaacac ccgcaagagc atcaccatcg gccccggccg cgccttctac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca ccctgaagca gatcgtgacc aagctgcagg ccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgggcag ccgagatcg tgatgcacag cttcaactgc 900
ggcgggcgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcgggccca acaacaccaa cggcaccatc accctgccct gccgcatcaa gcagatcacc 1020
aaccgctggg gcggaaggc catgtacgcc ccccccattcc gcggccagat ccgctgcagc 1080
agcaacatca ccggcctgct gctgaccgcg gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gccccggcgg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag 1200
tacaagggtg tgaagatcga gcccctgggc gtggccccc ccaaggccaa gcgcccgtg 1260
gtgcagcgcg agaagcgcg cgtgacctg ggcgccatgt tcttgggctt cctggggcgc 1320
gccggcagca ccatggggcg ccgcagcctg accctgaccg tgcaggcccg ccagctgctg 1380
agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tcgaggccca gcagcacctg 1440
ctgcagctga ccgtgtggg catcaagcag ctgcaggccc gcgtgctggc cgtggagcgc 1500
tacctgaagg accagcagct gctgggcattc tggggctgca gcggcaagct gatctgcacc 1560
accgccgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggg gcgcgagatc gacaactaca ccaacctgat ctacacctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tgggccagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tggcgggcct ggtgggcctg cgcacgtgtg tcaccgtgct gagcatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga cccgcttccc cgccccccgc 1920
ggccccgacc gccccgagg catcgaggag gaggcgggcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggacg acctgcgcag cctgtgcctg 2040
ttcagctacc accgcctgcg cgacctgatc ctgatcgccg ccgcatcgt ggagctgctg 2100
ggcgcccgcg gctgggaggc cctgaagtac tggggcaacc tgctgcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgccatcg ccacgcccgt ggccgagggc 2220
accgaccgca tcacgaggt ggcccagcgc atcgggcccg ccttctctgca catccccgcg 2280
cgcacccgcc agggcttcga gcgcgcctcg ctgtaactcg ag 2322

```

<210> 21

<211> 2310

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Lys121-Val200;
Asn425-Lys432

<400> 21

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttc tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
ccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggccctacgac 180
accgaggtgc acaacgtgtg ggccaccac gctgctgctg ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
ccggtgatca cccaggcctg cccaagggtg agcttcgagc ccaccccat ccactactgc 420
gcccccgccg gcttcgccat cctgaagtgc aacgacaaga agttcaacgg cagcgggccc 480
tgcaccaacg tgagcaccgt gcagtgcacc cacggcatcc gccccgtggg gagcaccag 540
ctgctgctga acggcagcct ggccgaggag ggctgtgtga tccgcagcga gaacttcacc 600
gacaacgcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcacccgc 660
cccaacaaca acaccgcaa gagcatcacc atcgccccg gccgcgctt ctacgccacc 720
ggcgacatca tcggcgacat ccgccaggcc cactgaaca tcagcggcga gaagtgaac 780
aacaccctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaagcaga gcagcgggcg cgaccccag atcgtgatgc acagcttcaa ctgcggcgcc 900
gagttcttct actgcaacag caccagctg ttcaacagca cctggaacaa caccatcggc 960
cccaacgaca ccaacggcac catcaccctg ccctgccgca tcaagcagat catcaacgcc 1020
cccaaggcca tgtacgcccc ccccatccgc ggccagatcc gctgcagcag caacatcacc 1080
ggcctgctgc tgaccgcgga cggcggaag gagatcagca acaccaccga gatcttccgc 1140
cccgcgggcg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtggtg 1200

```

```

aagatcgagc ccctgggagt ggccccccacc aaggccaagc gccgcgtggt gcagcgcgag 1260
aagcgcgcgc tgaccctggg cgccatgttc ctgggcttcc tgggcgcgcg cggcagcacc 1320
atgggcgccc gcagcctgac cctgaccgtg caggcccgcc agctgctgag cggcatcgtg 1380
cagcagcaga acaacctgct gcgcgccatc gaggcccgag agcacctgct gcagctgacc 1440
gtgtggggca tcaagcagct gcaggcccg gcgtggtgag tggagcgta cctgaaggac 1500
cagcagctgc tgggcatctg gggctgcagc ggcaagctga tctgcaccac cgccgtgccc 1560
tggaaaccca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gcgagatcga caactacacc aacctgatct acacctgat cgaggagagc 1680
cagaaccagc aggagaagaa cgagcaggag ctgctggagc tggacaagtg ggccagcctg 1740
tggaaactgt tcgacatcag caagtggctg tggtagatca agatcttcat catgatcgtg 1800
ggcggcctgg tgggcctgag catcgtgttc accgtgctga gcacgtgaa ccgcgtgcgc 1860
cagggttaca gccccctgag cttccagacc cgcttccccg cccccgcgg ccccgaccgc 1920
cccgagggca tcgaggagga gggcggcgag cgcgaccgag accgcagcag cccctggtg 1980
cacggcctgc tggccctgat ctgggacgac ctgcgagcgc tgtgcctgtt cagctaccac 2040
cgctgcgcgc acctgatcct gatcgccgcc cgcatcgtgg agctgctggg ccgccgcggc 2100
tgggaggccc tgaagtactg gggcaacctg ctgcagtact ggatccagga gctgaagaac 2160
agcgcctgta gcctgttcga cgccatcgcc atcgccgtgg ccgagggcac cgaccgcac 2220
atcgagggtg cccagcgcat cggccgcgcc ttctgcaca tccccgcgc catccgccag 2280
ggcttcgagc gcgccctgct gtaactcgag                                     2310

```

<210> 22

<211> 2298

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Ile201;
Ile424-Ala433

<400> 22

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcaa ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
catatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtagcac ccagctgctg 540
ctgaacggca gcctggccga ggaggcggtg gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcgcc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca ggccactgac aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcgcca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgcccctg cgcatcaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgcgg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
acccgcgacg gcggcaagga gatcagcaac accaccgaga tcttccgccc cggcgccggc 1140
gacatgcgag acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg cccccacaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgcctg 1260
acctggggcg ccatgttctt gggcttctct ggcgcgcggc gcagcaccat gggcgccgc 1320
agcctgaccc tgaccgtgca ggcccgccag ctgctgagcg gcacgtgca gcagcagaac 1380
aacctgctgc gcgccatcga ggcccgagc cactgctgac agctgaccgt gtggggcatc 1440
aagcagctgc agggccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740

```

```

gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagaccgg cttccccgcc ccccgcgccc ccgaccgccc cgagggcatc 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgctgttca gctaccaccg cctgcgcgac 2040
ctgacctga tcgccggccc catcgtggag ctgctggggc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgagggtggc 2220
cagcgcatcg gccgcgcctt cctgcacatc ccccgccgca tccgccaggg cttcgagcgc 2280
gccctgctgt aactcgag                                     2298

```

<210> 23

<211> 2298

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

Val120-Ile201B; Ile424-Ala433

<400> 23

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg tttcgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc ccaccgacc caacccccag 240
gagatcgatgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgcccggc 360
atcaccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggttcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtgagcac ccagctgctg 540
ctgaacggca gcctggccga ggagggcggt gtgatccgca gcgagaactt caccgacaac 600
gccaaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacacccc gcaagagcat caccatcggc cccggccgcg cttctacgc caccggcgac 720
atcatcggcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcgcca acaagaccat cgtgttcaag 840
cagcagcgc gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgccctgc cgcacaaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgagg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accgcgacg gcggcaagga gatcagcaac accaccgaga tcttcgccc cggcgggcgg 1140
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtgggtg agcgcgagaa gcgcgcccgtg 1260
accctgggcg ccatgttcct gggcttcctg ggcgcggccg gcagcaccat gggcgcccgc 1320
agcctgaccc tgaccgtgca ggcccgcag ctgctgagcg gcacgtgca gcagcagaac 1380
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1440
aagcagctgc agggccgctg gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgac tgcaccaccg ccgtgccctg gaacgccagc 1560
tgagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1740
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagaccgg cttccccgcc ccccgcgccc ccgaccgccc cgagggcatc 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgctgttca gctaccaccg cctgcgcgac 2040
ctgacctga tcgccggccc catcgtggag ctgctggggc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgagggtggc 2220
cagcgcatcg gccgcgcctt cctgcacatc ccccgccgca tccgccaggg cttcgagcgc 2280

```

gccctgctgt aactcgag

2298

<210> 24

<211> 2298

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val120-Thr202;
Ile424-Ala433

<400> 24

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgctg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggccctacgac 180
accgaggtgc acaacgtgtg ggccaccac gcttgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgggcggc 360
gccacccacg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gccggcttcg ccatactgaa gtgcaacgac aagaagttca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggc atccgccccg tggtagacac ccagctgctg 540
ctgaacggca gcctggccga ggagggcgct gtgatccgca gcgagaactt caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgccccaac 660
aacaacaccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcgcg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacaccat cggccccaac 960
aacaccaacg gcaccatcac cctgcccctg cgcataaagc agatcatcgg cggcgccatg 1020
tacgcccccc ccataccgcg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accgcgacg gcggcaaggc gatcagcaac accaccgaga tcttccgccc cggcgccgac 1140
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1200
ctgggcgtgg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgcccgtg 1260
accctgggcg ccatgttcct gggcttcctg ggcgcgcccg gcagcaccat gggcgcccgc 1320
agcctgaccc tgaccgtgca ggcccgcag ctgctgagcg gcacgtgca gcagcagaac 1380
aacctgctgc gcgccatcga ggcccagcag cactgctgac agctgaccgt gtggggcatc 1440
aagcagctgc agggccgctg gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1560
tgagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactgggtc 1740
gacatcagca agtggtgtg gtacatcaag atcttcatca tgatcgtggg cggcctgggtg 1800
ggcctgcgca tcgtgttcac cgtgctgagc atcttgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagacccg cttccccgcc ccccgcgccc ccgaccgccc cgagggcac 1920
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 1980
gccctgatct gggacgacct gcgcagcctg tgccgttca gctaccaccg cctgcgcgac 2040
ctgatactga tcgcgcgccc catcgtggag ctgctgggccc gccgcggctg ggaggccctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2160
ctgttcgacg ccatacgcct cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2220
cagcgcacgc gccgcgcctt cctgcacatc ccccgcgcca tccgccaggg cttcgagcgc 2280
gccctgctgt aactcgag

```

<210> 25

<211> 2358

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val127-Asn195

<400> 25

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgctgct ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgggggc agggaaactgc aacaccagcg tgatcaccca ggcctgcccc 420
aaggtgagct tcgagcccat ccccatccac tactgcgccc ccgccggctt cgccatcctg 480
aagtgaacg acaagaagtt caacggcagc ggcccctgca ccaacgtgag caccgtgcag 540
tgcaccacg gcatccgccc cgtgggtgagc acccagctgc tgctgaacgg cagcctggcc 600
gaggagggcg tgggtgatccg cagcgagaac ttaccgaca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc acccgcccc acaacaacac ccgcaagagc 720
atcaccatcg gccccggccg cgcttcttac gccaccggcg acatcatcgg cgacatccgc 780
caggcccact gcaacatcag cggcgagaag tggaaacaac ccctgaagca gatcgtgacc 840
aagctgcagg ccaggttcgg caacaagacc atcgtgttca agcagagcag cggcggcgag 900
cccgagatcg tgatgcacag cttcaactgc ggcggcgagt tcttctactg caacagcacc 960
cagctgttca acagcacctg gaacaacacc atcgccccca acaacaccaa cggcaccatc 1020
accctgccct gccgcatcaa gcagatcatc aaccgctggc aggaggtggg caaggccatg 1080
tacgcccccc ccatccgagg ccagatccgc tgcagcagca acatcacagg cctgtgtgtg 1140
accgcgcagc gcggcaagga gatcagcaac accaccgaga tcttcgccc cggcggcggc 1200
gacatgcgcg acaactggcg cagcgagctg tacaagtaca aggtggtgaa gatcgagccc 1260
ctgggctgtg cccccaccaa ggccaagcgc cgcgtggtgc agcgcgagaa gcgcgccgtg 1320
accctgggcg ccatgttcct gggcttctct ggcgcgcgcg gcagcaccat gggcgcgccg 1380
agcctgaccc tgaccgtgca ggcccgcag ctgctgagcg gcacgtgca gcagcagaac 1440
aacctgctgc gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc 1500
aagcagctgc agggccgcgt gctggccgtg gagegctacc tgaaggacca gcagctgctg 1560
ggcatctggg gctgcagcgg caagctgatc tgcaccaccg ccgtgccctg gaacgccagc 1620
tggagcaaca agagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1680
gagatcgaca actacaccaa cctgatctac accctgatcg aggagagcca gaaccagcag 1740
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1800
gacatcagca agtggctgtg gtacatcaag atcttcatca tgatcgtggg cggcctggtg 1860
ggcctgcgca tcgtgttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1920
cccctgagct tccagaccgg cttccccgcc cccgcgggcc ccgaccgccc cgagggcatc 1980
gaggaggagg gcggcgagcg cgaccgcgac cgcagcagcc ccctggtgca cggcctgctg 2040
gccctgatct gggacgacct gcgcagcctg tgctgttca gctaccaccg cctgcgcgac 2100
ctgatcctga tcgcccgcg catcgtggag ctgctggggc gccgcggctg ggaggccctg 2160
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag cgccgtgagc 2220
ctgttcgacg ccatcgccat cgccgtggcc gagggcaccg accgcatcat cgaggtggcc 2280
cagcgcacg gccgcgcctt cctgcacatc cccgcgcgca tccgccaggg cttcgagcgc 2340
gccctgctgt aactcgag                                     2358

```

<210> 26

<211> 2352

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Val127-Asn195;
Arg426-Gly431

<400> 26

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcg ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccacccac gcctgctgct ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gcgtgggggc agggaaactgc aacaccagcg tgatcaccca ggcctgcccc 420

```

```

aaggtgagct tcgagcccat ccccatccac tactgcgccc ccgcccggctt cgccatcctg 480
aagtgcacg acaagaagtt caacggcagc ggcccctgca ccaacgtgag caccgtgcag 540
tgcacccacg gcatccgccc cgtggtgagc acccagctgc tgctgaacgg cagcctggcc 600
gaggagggcg tggatgatcc cagcgagAAC ttcaccgaca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc acccgcccca acaacaacac ccgcaagagc 720
atcaccatcg gccccggccg cgccttctac gccaccggcg acatcatcgg cgacatccgc 780
caggccact gcaacatcag cggcgagaag tggaacaaca ccctgaagca gatcgtgacc 840
aagctgcagg ccaggttcgg caacaagacc atcgtgttca agcagagcag cggcggcgac 900
cccagatcg tgatgcacag cttcaactgc ggcgggcagt tcttctactg caacagcacc 960
cagctgttca acagcacctg gaacaacacc atcgggccca acaacaccaa cggcaccatc 1020
accctgccct gccgcacaa gcagatcatc aaccgcggcg gcggcaaggc catgtacgcc 1080
ccccccatcc gcggccagat ccgctgcagc agcaacatca ccggcctgct gctgaccgcg 1140
gacggcggca aggagatcag caacaccacc gagatcttcc gccccggggg cggcgacatg 1200
cgcgacaact ggcgagcga gctgtacaag tacaagggtg tgaagatcga gcccctgggc 1260
gtggcccca ccaaggccaa gcgcccgtg gtgcagcgcg agaagcgcgc cgtgaccctg 1320
ggcgccatgt tcctgggctt cctggggcgc gccggcagca ccatgggcgc ccgacgcctg 1380
accctgaccg tgcaggcccg ccagctgctg agcggcatcg tgcagcagca gaacaacctg 1440
ctgcgcgcca tcgaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag 1500
ctgcaggccc gcgtgctggc cgtggagcgc tacctgaagg accagcagct gctgggcac 1560
tggggctgca gcggcaagct gatctgcacc accgcctgct cctggaacgc cagctggagc 1620
aacaagagcc tggaccagat ctggaacaac atgacctgga tggagtggga gcgcgagatc 1680
gacaactaca ccaacctgat ctacaccctg atcgaggaga gccagaacca gcaggagaag 1740
aacgagcagg agctgctgga gctggacaag tgggcccagc tgtggaactg gttcgacatc 1800
agcaagtggc tgtggtacat caagatcttc atcatgatcg tgggcggcct ggtgggcctg 1860
cgcacgtgt tcaccgtgct gagcatcgtg aaccgcgtgc gccagggcta cagccccctg 1920
agcttccaga cccgttccc cgcggccgc ggcccgcacc gccccgaggg catcgaggag 1980
gagggcggcg agcgcgaccg cgaccgcagc agcccctgg tgcacggcct gctggccctg 2040
atctgggacg acctgcgcag cctgtgcctg ttcagctacc accgcctgcg cgacctgatc 2100
ctgatcgccg ccgcacatcg ggagctgctg ggccgcgcgc gctgggaggc cctgaagtac 2160
tggggcaacc tgctgcagta ctggatccag gagctgaaga acagcggcgt gagcctgttc 2220
gacgccatcg ccatcgccgt ggccgagggc accgaccgca tcatcgaggt ggccagcgc 2280
atcgccgcg ccttccctgca catccccgc cgcacccgc agggcttcga gcgcgcctg 2340
ctgtaactcg ag
2352

```